ELSEVIER
ACADEMIC
PRESS

Advances in PROTEIN CHEMISTRY

VOLUME 70

Fibrous Proteins: Coiled-Coils, Collagen and Elastomers

ADVANCES IN PROTEIN CHEMISTRY

EDITED BY

FREDERIC M. RICHARDS
Department of Molecular Biophysics
and Biochemistry,
Yale University
New Haven, Connecticut

DAVID S. EISENBERG
Department of Chemistry and Biochemistry
Center for Genomics and Proteomics
University of California, Los Angeles
Los Angeles, California

JOHN KURIYAN
Department of Molecular and Cellular Biology
University of California, Berkeley
Berkeley, California

VOLUME 70

Fibrous Proteins: Coiled-Coils, Collagen and Elastomers

EDITED BY

David A. D. Parry
Institute of Fundamental Sciences
Massey University
Palmerston North
New Zealand

John M. Squire
Biological Structure and Function Section
Biomedical Sciences Division
Imperial College London
London, United Kingdom

ELSEVIER
ACADEMIC
PRESS

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Volume 58
pp. 1-474 (2001)
Novel Cofactors
Volume 57
pp. 1-405 (2001)
Prion Proteins
Volume 56
pp. 1-567 (2001)
Drug Discovery and Design
Volume 55
pp. 1-438 (2001)
Evolutionary Protein Design
Volume 54
pp. 1-481 (2000)
Analysis of Amino Acid Sequences
Volume 53
pp. 1-442 (2000)
Protein folding mechanisms

Volume 70, Pages 1-525 (2005)

Previous vol/iss Next vol/iss

Fibrous Proteins: Coiled-Coils, Collagen and Elastomers
Edited by: David A. D. Parry, and John M. Squire
ISBN: 978-0-12-034270-9

Article List Full Abstracts
Add to my Quick Links
Display Selected Articles E-mail Articles Export Citations
articles **1 - 17**

1. **Series Editors**
Page iii
PDF (41 K) | View Related Articles

2. **Preface**
Pages xi-xii
SummaryPlus | Full Text + Links | PDF (27 K) | View Related Articles

3. **Fibrous Proteins: New Structural and Functional Aspects Revealed**
Pages 1-10
David A. D. Parry and John M. Squire
SummaryPlus | Full Text + Links | PDF (77 K) | View Related Articles

4. **Structural and Functional Implications of Sequence Repeats in Fibrous Proteins**
Pages 11-35
David A. D. Parry
SummaryPlus | Full Text + Links | PDF (496 K) | View Related Articles

5. **The Structure of Ŭ-Helical Coiled Coils**
Pages 37-38
Andrei N. Lupas and Markus Gruber
SummaryPlus | Full Text + Links | PDF (1842 K) | View Related Articles

6. **The Design of Coiled-Coil Structures and Assemblies**
Pages 79-112
Derek N. Woolfson
SummaryPlus | Full Text + Links | PDF (394 K) | View Related Articles

7. **Microdissection of The Sequence and Structure of Intermediate Filament Chains**
Pages 113-142
David A. D. Parry
SummaryPlus | Full Text + Links | PDF (492 K) | View Related Articles

8. **Intermediate Filament Associated Proteins**
Pages 143-202
Kathleen J. Green, Michael Bºhringer, Todd Gocken and Jonathan C. R. Jones
SummaryPlus | Full Text + Links | PDF (720 K) | View Related Articles

9. **Spectrin, Ŭ-Actinin, and Dystrophin**
Pages 203-246
M. J. F. Broderick and S. J. Winder
SummaryPlus | Full Text + Links | PDF (448 K) | View Related Articles

10. **Fibrinogen and Fibrin**
Pages 247-299
John W. Weisel
SummaryPlus | Full Text + Links | PDF (694 K) | View Related Articles

11. **Molecular Structure of the Collagen Triple Helix**
Pages 301-339
Barbara Brodsky and Anton V. Persikov
SummaryPlus | Full Text + Links | PDF (776 K) | View Related Articles

12. **Collagen Fibril Form and Function**
Pages 341-374
T. J. Wess
SummaryPlus | Full Text + Links | PDF (704 K) | View Related Articles

13. **Molecular Packing in Network-Forming Collagens**
Pages 375-403
Carlo Knupp and John M. Squire
SummaryPlus | Full Text + Links | PDF (783 K) | View Related Articles

14. **Fibrillin Microfibrils**
Pages 405-436
Cay M. Kielty, Michael. J. Sherratt, Andrew Marson and Clair Baldock
SummaryPlus | Full Text + Links | PDF (681 K) | View Related Articles

15. **Elastin**
Pages 437-461
Suzanne M. Mithieux and Anthony S. Weiss
SummaryPlus | Full Text + Links | PDF (323 K) | View Related Articles

16. **Author Index**
Pages 463-506
PDF (162 K) | View Related Articles

17. **Subject Index**
Pages 507-525
PDF (96 K) | View Related Articles

articles **1 - 17**

Browse Search My Settings Alerts Help
Home

About ScienceDirect | Contact Us | Terms & Conditions | Privacy Policy

Copyright É 2008 Elsevier B.V. All rights reserved. ScienceDirectÈ is a registered trademark of Elsevier B. V.

PREFACE

Fibrous proteins represent a substantial subset of the human proteome. They include the filamentous structures found in animal hair that act as a protective and thermoregulatory outer material. They are responsible for specifying much of an animal's skeleton, and connective tissues such as tendon, skin, bone, cornea and cartilage all play an important role in this regard. Fibrous proteins are frequently crucial in locomotion and are epitomised by the muscle proteins myosin and tropomyosin and by elastic structures like titin. Yet again the fibrous proteins include filamentous assemblies, such as actin filaments and microtubules, where these provide supporting structures and tracks for the action of a variety of molecular motors.

It is nearly 20 years since this field was fully reviewed and there have been very significant advances in that time. The present book, therefore, represents one of a set of three volumes in the Elsevier '*Advances in Protein Chemistry*' series covering the entire fibrous protein field. These are entitled:

Fibrous Proteins: Coiled-Coils, Collagen and Elastomers
Fibrous Proteins: Muscle and Molecular Motors
Fibrous Proteins: Amyloids, Prions and β-Proteins

The present volume covers 'Coiled-Coils, Collagen and Elastomers'. The first few Chapters describe the importance of α-helical coiled-coil proteins with examples of the intermediate filament proteins, the spectrin superfamily and fibrin/fibrinogen. It is shown that the design principles governing the structures of coiled-coil proteins are largely discernible and can be specified with a high degree of confidence, thanks in large part to the wealth of crystal structures now available. Within the connective tissues covered here, constituents of defining importance mechanically are collagen fibrils and networks, and elastic fibres. Details are given of crystal structures of collagen peptides and the effects on conformation of the precise sequence of the distinct constituent triplets. The ultrastructures of connective tissues are largely defined by the spatial arrangement of the collagen fibrils and networks, and these too are elucidated in some detail. The final part of the book covers elastic fibres with their elastin cores and fibrillin-containing microfibril palisades. A common theme throughout is the increased characterisation of the structures and functions of mutants. Some of these occur naturally and lead to disease, whilst others have been genetically engineered in order to study design principles.

The complete set of three books will provide a compendium of up-to-the-minute information on the entire fibrous protein field. Each Chapter, which is clearly written, fully illustrated and with a comprehensive citation list, is by an acknowledged authority in the field. It is our hope that, together, these books will enable valuable comparisons to be made, they will allow general principles to be elucidated and they will help to take the fibrous protein field forward in good shape into the 21st Century era of post-genomics, molecular medicine and nanoscience.

David Parry
Massey University
New Zealand

John Squire
Imperial College London
United Kingdom

FIBROUS PROTEINS: NEW STRUCTURAL AND FUNCTIONAL ASPECTS REVEALED

By DAVID A. D. PARRY* AND JOHN M. SQUIRE†

*Institute of Fundamental Sciences, Massey University, Palmerston North 5301, New Zealand;
†Biological Structure and Function Section, Biomedical Sciences Division, Imperial College London, London SW7 2AZ, United Kingdom

I. Introduction 2
II. α-Fibrous Proteins 2
III. Connective Tissue 6
IV. Summary 9

Abstract

Coiled-coil proteins, collagen, and elastomers together comprise an important subset of the fibrous proteins. The former group—the α-fibrous coiled-coil proteins—are widely distributed in nature and, indeed, the characteristic heptad motif has been recognized as an oligomerisation motif in fibril-forming collagens. This volume has selected a number of the α-fibrous proteins for detailed discussion, including intermediate filament proteins, the spectrin superfamily, and fibrin/fibrinogen. Of particular interest is the growing realization that the design principles governing the structures of these coiled-coil proteins are now largely discernible and can be specified with a high degree of confidence, due in large part to the wealth of crystal structure data now available. Within the connective tissues covered in this volume, two constituents of defining importance mechanically are the collagen fibrils/networks and the elastic fibers. Crystal structures of collagen peptides have been published and are described. The effects of the precise sequence of the distinct constituent triplets on molecular conformation have also become clearer. The ultrastructures of connective tissues are largely defined by the spatial arrangement of the collagen fibrils and networks, and this is elucidated here in some detail. The elastic fibers with their elastin cores and fibrillin-containing microfibril palisades are also described. A theme underlying all of the proteins discussed in this volume is the significantly increased effort to characterize the structures and functions of mutants. Some of these occur naturally and lead to various disease states, while others have been genetically engineered in order to study design principles.

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70001-7

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

I. Introduction

Over the past decade, rapid progress has been made in gaining an understanding of the structure and function of fibrous proteins and it is now timely to produce a comprehensive treatise on the subject. Towards that end, three volumes of Advances in Protein Chemistry have been commissioned by Elsevier on the general theme of ''Fibrous Proteins and Related Structures.'' The first volume focused on ''Molecular Motors and Muscle'' and was edited by Squire and Parry. This volume, the second in the series, examines ''Coiled Coils, Collagen, and Elastomers'' (edited by Parry and Squire). The third volume, edited by Kajava, Squire, and Parry, discusses ''Amyloids, Prions, and β-Proteins.'' Together, these volumes cover most of the key developments that have occurred in recent years and a wealth of information is provided for the reader by experts in the field. It is hoped that these three volumes will provide a solid and up-to-date basis for further studies and will help to identify areas requiring new insights.

II. α-Fibrous Proteins

Fibrous protein sequences are often characterized by the presence of simple repetitive motifs. Some are exact in length and/or sequence, but others are only approximate and display considerable variation. Some motifs contain residues that are absolutely conserved in some positions, whereas in others it is only the sequence character that is maintained over the repeat length. In many fibrous proteins the repeats occur contiguously, whereas in others they are found widely separated in the sequence. The varieties of sequence repeat that have been observed are typed and catalogued here by Parry (Chapter 2). Each motif forms a discrete element of structure; in many instances, these are arranged helically with respect to one another. In many cases an elongate structure is formed, and this can lead naturally to molecular aggregation and the formation of functional filaments.

The structures of many of these motifs have already been characterized by X-ray crystallography and nuclear magnetic resonance (NMR) methods. A particular motif of note is the underlying heptad (and related nonheptad) substructure present in α-fibrous proteins. The design principles of this class of structure are becoming increasingly well understood in terms of amino acid sequence. Furthermore, the importance of the interactions that arise between and within chains to stabilize and specify the secondary and tertiary structure has become better appreciated. Much of this progress can be attributed to the detailed crystallographic studies

that have been undertaken using fragments of native proteins and specifically designed variants. Lupas (Chapter 3) summarizes the latest structural information on α-fibrous proteins and, in so doing, reveals the stunning progress that has been made in recent years. The wide variation in conformation exhibited in the crystal structures has challenged what we really mean by the term *coiled coil*, and it is clear that the spectrum of structures extends well beyond that which was initially considered ''standard.'' As Lupas notes, we now have fibers, zippers, tubes, sheets, spirals, funnels, and rings that are all specified by coiled-coil packing arrangements. Computer analyses of the amino acid sequences of these proteins have begun to provide rules for distinguishing one multimer from another, but a challenge still exists in recognizing imperfect coiled coil repeats in a sequence. Indeed, breaks in these underlying substructures (stutters, stammers, skips) have made recognition a much more difficult exercise than would have been apparent after the first sequence of an α-fibrous protein (tropomyosin) was derived some 30 years ago, This, of course, displayed a perfect heptad repeat, but tropomyosin is now seen as an exception to the rule. The physical significance of stutters and skips (i.e., heptad and nonheptad phase discontinuities) is reasonably well understood in terms of local unwinding of the constituent α-helical strands. However, no crystal structure of a stammer region has yet been determined.

Woolfson's review (Chapter 4) neatly complements that of Lupas. He describes the real progress that has been made on the design principles underlying the coiled-coil conformation. In particular, he addresses the rationale behind the presence of certain residues or residue types in specific positions in the underlying heptad substructure characteristic of many α-fibrous proteins. These features drive the assembly of the constituent right-handed α-helices towards a left-handed multistranded structure, known here as a canonical coiled coil. However, the heptad repeat is not the only one of importance that can lead to a coiled-coil structure. Indeed, nonheptad repeats may result in noncanonical coiled coils that can lack either left-handed or regular geometry. For example, right-handed coiled coils can be formed from hendecad (11 residue) sequence repeats. The distribution of charged and apolar residues gives rise to amphipathic α-helices for both heptad and nonheptad repeats, but subtleties in the presence of branched or unbranched apolar residues in positions *a* and *d* (for instance) and appropriately charged residues in positions *e* and *g* (or sometimes in positions *b* and *c*, also) determine the appropriate number of strands, the formation of homo- or heteromultimers, their relative axial alignment, and their relative polarity. The coiled coil has proved to be an ideal system to explore the role(s) played by individual amino acids in determining both structure and function.

Woolfson's review provides a summary of the state of the art in coiled-coil design, and provides a keen insight into both the future potential of design methods and the challenges that they face.

Intermediate filament proteins, and others intimately associated with them, have become of major interest to cell biologists, especially in the past two decades. However, the history of research on wool and the hard α-keratins goes back much further, to the mid 1930s. This field was small, very specialized, and not well understood by the wider scientific community. This situation changed dramatically soon after the discovery of a third set of filaments in cells that were intermediate in size between the large microtubules and the small microfilaments. It was recognized that the wool and hair keratins were closely related by both sequence and structure to the intermediate filaments. This was a defining moment in the field. Almost overnight, the detailed structural work available on the hard α-keratins was related to the chemical data that were so easily gained (in comparison to the hard α-keratins) on the intermediate filaments. This understanding revolutionized the field and led to many subsequent significant advances.

Parry (Chapter 5) describes the structure of the intermediate filament proteins seen in hair, skin, vimentin, and neurofilaments (among others). Here he details the data that allow the role of individual residues and short lengths of sequences in these proteins to be understood in terms of secondary and tertiary structure, aggregation characteristics, and function. Site-directed mutagenesis studies have been particularly important in this regard, as has a comparison of the large number of intermediate filament sequences now available. These studies have enabled potentially conserved inter- and intra-ionic interactions to be recognized, as well as specifying those residues involved in stabilizing the four particular modes of molecular aggregation characterized from crosslinking data.

Intermediate filament associated proteins (IFAPs) are also key components of the system. Green, Jones, and colleagues (Chapter 6) provide a cogent account of the structure–function relationships of a number of these, including filaggrin, trichohyalin, and increasingly numerous members of the plakin family. A host of other IFAPs have also been characterized in recent years, and these too are discussed in detail. It has long been known, of course, that IFAPs have a critical role in linking together elements of the intermediate filament scaffold in cells. What has been less well understood, but which is now made very evident from the review of Green *et al.*, is that the IFAPs have a large range of functions far beyond those relating purely to structural connections. Indeed, Green *et al.* show that IFAPs can be molecular motors, chaperones, enzymes, adaptors, membrane receptors, or proteins that relate to cell homeostasis and

metabolism. IFAPs, originally the poor relation to the IF proteins, have now very clearly come into their own. It is only by understanding the interactions of the IF–IFAP complex that a full understanding can be gained of this complex group of biological tissues that include, among many others, skin and hair.

Triple-helical α–helical coiled coils are achieved *in vivo* in two very distinct ways. In the first, a single chain folds back on itself to form a structure with two similarly and one oppositely directed ''strand.'' These coiled coils are exemplified by the (related) structures seen in the rod domain of members of the spectrin superfamily of α-spectrin, α-actinin, and dystrophin/utrophin, each of which is derived from a common α-actinin ancestor (Broderick and Winder, Chapter 7). The dumbbell-shaped molecules have a flexible rodlike structure with membrane-anchoring capabilities. This superfamily is specifically characterized by the presence of contiguous spectrin repeats, each of which forms the three-α–helix motif noted above. Actin-binding domains (N-terminal domain) are found in all members except α-spectrin; EF hands (C-terminal domain) are found in all members except β-spectrin. Some members of the superfamily also contain additional domains specific to their particular function, thus extending the group's role considerably beyond that of simply crosslinking actin filaments together in order to form bundles. As with many other fibrous proteins, considerable efforts have been made in recent years to understand the effects produced by mutations observed in various disease states, not least of which (in dystrophin) relates to muscular dystrophy. The three-dimensional conformations of actin-binding domains, for example, exhibited in most members of the spectrin superfamily have now been determined at atomic resolution by either X-ray crystallography or NMR studies. These have been very informative in identifying residues that have special importance structurally and functionally.

The second (and more common) situation of a triple-helical α–helical coiled coil occurs when three chains aggregate, usually parallel to one another and with the chains in axial register. An example in this case is provided by the blood-clotting protein fibrin/fibrinogen. Marked progress on the structure and function of fibrin/fibrinogen has been made in recent times and, as a result, Weisel's review of the field (Chapter 8) will be seen as particularly timely to the research community. Structurally, X-ray diffraction studies on a variety of fibrinogen fragments have revealed details of the molecular conformation, which consists essentially of a central region linked via triple-stranded coiled-coil domains to globular domains lying at each end of the 45-nm-long molecule. Consequentially, chain connectivity is now well understood, as are the disulphide bond

connections and the domain organization. This picture has assisted in advancing our functional knowledge of fibrinogen, especially with regard to its metabolism and biosynthesis. Weisel reports that research is now very close to revealing the molecular mechanisms of fibrin polymerization. Fibrinogen/fibrin represents a rapidly advancing field, which has proved both exciting for those involved and rewarding in terms of our increased understanding of one of nature's most important molecules.

III. Connective Tissue

Connective tissues are a family of closely related but structurally/functionally diverse biological materials. Their importance in defining the skeleton, our bodies, and our mode of locomotion is beyond dispute. Likewise, the mechanical attributes of connective tissues in being able to withstand external and internal stresses, and in providing integrity to almost every tissue in the body, are unequivocally of great importance in sustaining life as we know it. This volume is concerned with disseminating the latest knowledge on crucial components of connective tissues—collagen (Chapters 9–11) and elastic fibers (Chapters 12 and 13).

It rarely fails to astound those working in the field of connective tissue that nature has been able to produce such a diverse range of mechanical and functional materials using essentially the same components—collagens, proteoglycans, elastic fibers, minerals, water, cells, and a variety of rather more minor proteins. Admittedly, these components are used in very different proportions in various tissues, and genetically distinct types of collagen as well as diverse proteoglycans (for example) are incorporated where appropriate. The three-dimensional organization of these elements is also highly specialized and this has evolved to allow the function of each tissue to be optimized. In some cases, the collagen molecules aggregate to form fibrils, sometimes with a very narrow distribution of sizes (as in the cornea), but other times with a wide range of diameters (as in tendons). In other cases, the collagen molecules form distinct and well characterized networks, basement membrane being one of particular note. For largely functional reasons, collagens are increasingly found in heteropolymeric rather than homopolymeric structures. These heteropolymeric fibrils and/or networks interact in a highly precise manner with specific proteoglycans to specify superaggregates that determine many of the mechanical attributes displayed by the intact tissue. Collagen fibrils assemble to form even larger ensembles, and they do so in a manner directly related to their function. For example, the fibrils can form unidirectional fibers (as in tendons), two-dimensional layers (as in skin), or more complex three-dimensional arrangements (as in cartilage and the cornea). Elastic fibers

are present in virtually all connective tissues, though the amount present varies greatly. These fibers are present most often (and in the highest quantities) in tissues that are frequently flexed, as in the necks (ligamentum nuchae) of grazing animals. The variation and general decrease in water content with age also impinges directly on the mechanical attributes of the tissue as a whole.

The classical concept that collagen molecules all have an extensive region in their amino acid sequences with an uninterrupted triplet substructure of the form $(G\text{-}X\text{-}Y)_n$ has not been tenable for many years. Indeed, the collagen types that have now been characterized (27 at the latest count) vary significantly from one another in this regard. For example, fibril-forming collagens tend to show a very long and continuous triplet substructure (338–341 triplets). In contrast, nematocyst minicollagens contain only 14 triplets. Types IV and VII collagen are different again, as they contain multiple (more than 20) breaks in the phasing of their triplet substructures. These have the effect of enhancing molecular flexibility. In the case of C1q, a single discontinuity results in a rigid kink being formed (Brodsky and Persikov, Chapter 9). Thus, there is considerable variation on the triplet theme, not only in the collagens but also in many other proteins containing triplet subdomains.

Crystal structure data on collagen peptides have revealed exciting information on the triple-helical structure and the manner by which it is stabilized through hydrogen bonding networks and water interactions. Interestingly, the helical parameters have been shown to be sequence-dependent with gly–pro–hyp sequences having 7/2 symmetry and more general sequences being more akin to 10/3. Allied to the latest data on collagen mutations, Brodsky and Persikov provide a comprehensive contribution on collagen structure and function that brings the field right up to date.

Wess (Chapter 10) considers the structure of fibril-forming collagens. These generally constitute the bulk of the collagen present in a tissue; they have particular importance through their size and ultrastructural organization in providing considerable mechanical strength to the tissue. The fibrils have an axial period of about 65–67 nm and diameters ranging up to about 500 nm. In all cases, the lateral dimensions are highly regulated, especially so for the uniform fibril diameters seen in the cornea, a factor related directly to transparency. Frequently, more than one collagen type is found in these fibrils. Sometimes they appear predominantly in the core of the fibril, but in other cases they are found on the fibril surface, where they can interact most easily with other molecules to provide the necessary functionality. FACIT collagen (Fibril-Associated Collagens with Interrupted Triple helices) is the terminology used to describe the latter type

of molecules that bind specifically at a fibril surface. Examples include collagen types IX, XII, and XIV. The packing of collagen molecules has been determined using X-ray diffraction techniques. From these data it is possible to assess the manner in which these molecules group together to form subfibrillar elements, such as microfibrils. This is also described by Wess in Chapter 10.

Other collagens, such as types IV, VI, VIII, and X, do not form fibrils. Instead, they aggregate to form meshes and open networks, thereby providing a selective molecular filter/barrier or a supporting structure for the tissue. The N- and C-terminal regions in these molecules lack a triplet substructure, but often play an important role in facilitating network formation through interactions between constituent molecules. Because they are often intrinsically disordered, elucidating details of the structure of this type of network is difficult (sometimes extremely so), but Knupp and Squire (Chapter 11) provide a comprehensive analysis of the pertinent data relating to the most common network-forming collagens. They also describe a possible collagen ''segmented supercoil'' motif that, in addition to the N- and C-terminal interactions, may be an important mode of intermolecular interaction. Other features of similarity and difference in network-forming collagens are highlighted.

In order to understand more of the complex interrelationships defining the structure and function of connective tissues, it is necessary to consider not only collagen but also other components. In this volume, however, discussion has been confined to the role played by the elastic fibers. These are largely (but not totally) composed of two proteins: fibrillin molecules that assemble to form a circumferentially arranged picket fence of microfibrils, and a highly crosslinked mass of tropoelastin molecules that constitute the central core. Specifically, it is fibrillin-1 that is the major constituent of the microfibrillar component of elastic fibers. This glycoprotein is about 160 nm in length and consists of 47 EGF-like domains, seven TB domains, two hybrid motifs, and a putative hinge region rich in proline residues. At present, there is much debate on the precise arrangement of the fibrillin-1 molecules in the 10- to 12-nm diameter microfibrils and their mechanical attributes; Kielty *et al.* (Chapter 12) discuss the ''hinged'' and ''staggered'' models in some detail. Microfibril-associated proteins, of which an increasing number have now been identified, are also significant both in the assembly process and in defining the function of the microfibril *in vivo*. While both aspects are currently imperfectly understood, it is clear that progress is being made. The microfibrils, assembled under a partially cell-regulated process, provide an outer framework within which the deposition of tropoelastin molecules is directed, but it is relevant to note here that microfibrils are also present in substantial

numbers in tissues that lack elastin. Mutations in fibrillin-1 have also been characterized and shown to cause Marfan syndrome, a heritable disease resulting in severe aortic, ocular, and skeletal defects.

Tropoelastin, on the other hand, is a soluble, medium-size protein (about 60–70 kDa) that becomes intimately attached to other tropoelastin molecules through lysine-mediated crosslinks. Elastin, as it is then termed, constitutes the bulk of the elastic fibers (about 90%) and is a resilient, elastic, but insoluble assembly with an extraordinarily long half-life (70 years). As noted earlier, elastin interacts with and is surrounded by the fibrillin-containing microfibrils and their associated proteins, thereby generating an entity *in vivo* that is perfectly designed for its biomechanical role *in vivo*. Debate in the elastin field centers on both the structural form adopted by elastin and the mechanism by which elasticity is achieved. Mithieux and Weiss (Chapter 13) discuss and compare the four major conformational models that have been proposed—random chain, liquid drop, oiled coil, and fibrillar. Each is structurally distinct, but all of them invoke entropic changes as the driving force (albeit in slightly different ways) to explain the mechanism by which elasticity is conferred to the elastic fibers.

The discussion here focuses on the structure and function of individual components of connective tissue (collagens, elastin, and fibrillin) rather than trying to integrate these components into the functioning whole. This latter goal, of course, represents the Holy Grail for those working in the field, but it is only by understanding the components in greater detail that we are likely to comprehend and appreciate the working of the entire system of interacting elements.

IV. Summary

Fibrous proteins provide a number of challenges to those seeking to understand them in detail at the molecular level. Not least of these challenges is trying to crystallize fibrous proteins in a form suitable for structural investigation using X-ray diffraction techniques. Even crystallization of fragments has proved problematic, though real progress has now been made and some informative results have been obtained for intermediate filament proteins, muscle proteins, fibrinogen, collagen, and many other fibrous molecules. In some cases, NMR methods have allowed the crystallization step to be avoided, thus permitting proteins to be studied in solution. Of course, NMR methods are not without their particular problems and most structural data have been obtained using X-ray diffraction.

In spite of these difficulties (or perhaps because of them), a suite of both chemical and physical methods has been devised to facilitate the research community in its quest to gain an in-depth understanding of the fibrous proteins. Some of these are theoretical and some experimental. Among the former methods are those based on bioinformatics. These methods use modeling and pattern recognition techniques to identify sequence and structural motifs previously discovered and characterized in detail. These motifs are often short in length and repeated consecutively many times. Indeed, these two features are characteristic of fibrous proteins (Chapter 2). Repeats are not confined to fibrous proteins and many are found in globular proteins (though these are usually quite long in comparison to those seen in the fibrous proteins). Consequently, they lead to domains of reasonable size (perhaps 20–50 residues but sometimes much larger). An appreciable number of α-fibrous structures have now been solved at atomic resolution (Chapter 3), allowing some of the key structural principles to be recognized and incorporated in *de novo* design methods (Chapter 4). Much progress has also been made on specific members of the class of α-fibrous structures. These include the two-stranded (double-helical) intermediate filament molecules (Chapter 5) and some of their associated proteins (Chapter 6), the single-stranded (but triple-helical) spectrin/α-actinin/dystrophin molecules (Chapter 7), and also the three-stranded (triple-helical) fibrin/fibrinogen molecule (Chapter 8).

Connective tissues, as noted earlier, are composite materials par excellence. Before the entire tissue can be appreciated in its full glory, it is necessary to probe details of the structure, function, and role of each component. This has been done here for collagen (Chapter 9), where the nature of the underlying triplet structure has been systematically investigated and crystal structures reported. Chapter 10 deals with collagen molecules that are fibril-forming, whereas Chapter 11 concentrates on those that form networks. Both forms of aggregation are functionally important, although the emphasis mechanically is clearly quite different in the two cases. Elastic fibers, defined by an outer palisade of fibrillin-containing microfibrils and an inner core of tropoelastin, are described in Chapters 12 and 13. These fibers are rarely accorded the importance that they merit, but the contributions in this volume will surely do them the justice they deserve.

STRUCTURAL AND FUNCTIONAL IMPLICATIONS OF SEQUENCE REPEATS IN FIBROUS PROTEINS

By DAVID A. D. PARRY

Institute of Fundamental Sciences, Massey University, Palmerston North 5301, New Zealand

I. Introduction 12
II. Sequence Regularities 13
A. Fixed Length and Exact Sequence Repeats (Type A Repeat) 13
B. Fixed Length Repeats but Residues Absolutely Conserved in Only Some Positions of the Motif (Type B Repeat) 13
C. Fixed Length Repeats with Sequence Character Maintained (Type C Repeat) 14
D. Nonintegral Repeats in Sequence Character (Type D Repeat) 14
E. Sequence Motifs of Variable Length in Proteins (Type E Repeat) 14
III. Effect of Sequence Repeats on Secondary Structure, and Implications for Tertiary and Quaternary Structure 15
A. Type A Repeats 15
B. Type B Repeats 17
C. Type C Repeats 19
D. Type D Repeats 23
E. Type E Repeats 27
IV. Summary 30
References 32

Abstract

The amino acid sequences of increasingly large proteins have been determined in recent years, and it has become more and more apparent that within these sequences nature has employed only a finite number of structural/functional motifs. These may be strung along the sequence in tandem and, in some cases, several hundred times. In other instances, the positions of the motifs show little obvious order as regards to their relative linear arrangement within the sequence. The observed sequence repeats have been shown to vary in size over at least two orders of magnitude. It is shown here that the repeats can readily be classified on the basis of character, and five distinct groups have been identified. The first of these (Type A) represents those motifs that are fixed in length and conserved absolutely in sequence (>99%); the second (Type B) includes motifs that are also fixed in length, but where absolute sequence conservation occurs only in some positions of the repeat. The third category (Type C) contains fixed length motifs, but the character of only some of the positions in the motif is maintained. The fourth group (Type D) includes motifs that have

DOI: 10.1016/S0065-3233(04)70002-9
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

nonintegral lengths. The fifth class (Type E) contains motifs, often displaying some variations in their lengths even within a single species, which maintain a discrete structural form related directly to their function. Examples are presented for each category of repeat, and these are drawn almost exclusively from the fibrous proteins and those proteins that are normally associated with them *in vivo*.

I. Introduction

Fibrous proteins are often designed specifically to aggregate and form filamentous assemblies. Examples of these include fibrous proteins in the collagen and α-fibrous classes. Filament formation requires that complementary groups in two or more molecules lie in appropriate axial and azimuthal orientations to facilitate interaction. For example, the presence of periodic clusters of acidic residues in one molecule might lie close to periodic clusters of basic residues in another. This could facilitate the formation of a network of stabilizing intermolecular ionic interactions and thereby specify a unique mode of aggregation. Likewise, patches of apolar residues in different molecules might lie in positions where the two areas could come together and shield both regions from the aqueous environment, thereby stabilizing and specifying the mode of assembly. Complementarity of shape and hydrogen-bonding potential are other means used *in vivo* to provide specificity of interaction and assembly. However, in order for these possibilities to be realized for filament-favoring molecules, two special sequence-related features are required. First and by definition, a filamentous structure implies the presence of a relatively short-range motif repeated contiguously. Such a motif will adopt a particular conformation, often one of the well-known elements of secondary structure (α-helical, β-strand, or collagen-like). Since such residues or clumps of residues would naturally favor similar environments, it is extremely likely that the elements will be related to one another helically, and that an elongate molecular structure will be formed. Secondly, a regular pattern of intermolecular interactions is needed for filamentous assembly to occur; this also infers a corresponding regularity in the underlying amino acid sequences of the interacting molecules.

Sequence repeats in proteins vary greatly. Some are very short and others are extremely long. The repeats can be exact or approximate, they can contain residues that are absolutely conserved in some positions but not in others, and they can be extremely imprecise in other than a conserved general character. Some repeats are fixed in length and others are not. Some are primarily functional and, in an apparent contradiction in terms, can also vary in length. Some repeats occur many times

contiguously while others are found only a small number of times consecutively. Still others are found distributed singly in what appears to be (but is clearly not) a random manner along the entire length of the protein chain. Many motifs are now well established and these have been recognized in quite diverse proteins. The PROSITE database gives a summary of these motifs (http://www.expasy.org/prosite/) and, as such, provides the researcher with a first indication of the repeats that exist in a newly determined protein sequence (Hulo *et al.*, 2004; Sigrist *et al.*, 2002).

In order to bring some order to what is clearly a diverse array of observations, this review has attempted to categorize the repeats. Five such classes have been recognized. Each will be dealt with in turn and appropriate examples presented. No attempt has been made to list the repeats observed in the sequences of all proteins; instead, representative examples are presented from the fibrous proteins in particular, and the proteins associated with them, in order to illustrate key features of some of the more interesting structures. First, however, these classes are defined.

II. Sequence Regularities

A. *Fixed Length and Exact Sequence Repeats (Type A Repeat)*

Type A repeats contain exact numbers of residues (described here as quantal). Exact copies of the repeat are often, though not invariably, strung along the sequence contiguously. The extent of the motif varies over several orders of magnitude and ranges from two residues in some silks (Parry, 1979) to over 500 residues in epiplakin (Fujiwara *et al.*, 2001). In the latter case, the repeat (occurring five times in tandem) is conserved at the 99.6% level. Any repeat conserved at the 99% level or above is considered here to be exact. Motifs longer than those in epiplakin will undoubtedly be found, and these too will arise from gene duplication events.

B. *Fixed Length Repeats but Residues Absolutely Conserved in Only Some Positions of the Motif (Type B Repeat)*

The classic example of a Type B repeat is that presented by the α-chains in collagen. Three such chains aggregate to form a triple-helical collagen molecule, but they can only do so if glycine is positioned in every third residue of the sequence (Hulmes, 1992). This is because glycine is located internally and, due to its size, is the only residue that can fit stereochemically into the space available. In the Type I collagen α-chain, the repeat occurs 338 times contiguously.

C. *Fixed Length Repeats with Sequence Character Maintained (Type C Repeat)*

Type C repeats are very common in proteins. They are quantal in length, but the repeats themselves do not contain residues that are conserved absolutely in any position. However, several positions within the repeats are strongly conserved in character. A classic example of a Type C repeat is that given by the heptad substructure in α-fibrous proteins. This has the form $(a\text{–}b\text{–}c\text{–}d\text{–}e\text{–}f\text{–}g)_n$ with the *a* and *d* positions generally occupied by apolar residues, and the *e* and *g* positions by charged or hydrophilic residues. The heptad is characteristic of an α-helical conformation (Cohen and Parry, 1986, 1990; Lupas, 1996), but comparison of any two sequences with a heptad substructure generally reveals only about 15–20% identity. The motif also implies that several α-helices will aggregate to form a multistranded left-handed coiled-coil rope to shield the apolar stripes on the surface of the α-helices from the aqueous environment.

D. *Nonintegral Repeats in Sequence Character (Type D Repeat)*

For many proteins, it is the geometry of the molecules that is paramount; packing considerations are secondary. Consequently, there may be nonquantal periodicities of residues and, more often, differences in residue types on the surface of the molecule. These are important in specifying molecular interactions and assembly. Examples of the Type D repeat are formed for both the acidic residues and the basic residues in two major coiled-coil segments of keratin (McLachlan, 1978; Parry *et al.*, 1977) and desmin/vimentin intermediate filament chains (McLachlan and Stewart, 1982). Provided that the interacting molecules display one of the possible relative axial staggers, they will necessarily be stabilized in the filament by clusters of intermolecular ionic interactions. Other examples are that seen for tropomyosin, which has a regular distribution of acidic residues and, to a lesser extent, apolar residues (McLachlan and Stewart, 1976; Parry, 1975) that match the period exhibited by the actin monomers in the thin filaments of muscle. This allows each actin monomer to be regulated by tropomyosin in a similar manner.

E. *Sequence Motifs of Variable Length in Proteins (Type E Repeat)*

A sequence repeat found in both fibrous and globular proteins is one in which a functional motif occurs, possibly several times in the sequence but often noncontiguously. This Type E repeat is exemplified by the Ca^{2+} EF hand, which is a length of sequence specifying a pair of α-helices that

define a site capable of binding a Ca^{2+} (or sometimes a Mg^{2+}) ion. Some residues within this repeat are largely, but not absolutely, conserved while other positions show a wide range of tolerance to different residues. The absolute lengths of some Type E repeats vary, with small deletions or insertions possible without compromising function.

III. Effect of Sequence Repeats on Secondary Structure, and Implications for Tertiary and Quaternary Structure

A. *Type A Repeats*

Short exact repeats of sequence generally adopt a well-defined element of secondary structure. If these repeats in the protein are arranged in tandem, then a helical arrangement is likely since this will result in similar spatial environments for each motif. A fibrous structure with a high axial ratio is the natural result and the repeat will primarily be a structural one. In the situations where the exact repeats are not consecutive, it is more probable that the motifs will be recognition or interaction sites, and hence have more of a functional than structural significance. On the other hand, long exact repeats lacking internal sequence substructure are likely to form discrete domain structures (possibly globular). Thus, the longer the repeat length, the more likely it is that the protein can be represented by a string of globular domains. It is unlikely in these circumstances that the motifs will be related to one another helically.

Silks that contain short sequence repeats, often incorporating only a small subset of the available amino acids, seem relatively common. For example, *Pachylota audouinii* contains equal amounts of alanine and glutamine residues, and this silk is thus believed to have a $(A\text{-}Q)_n$ substructure (Lucas and Rudall, 1968). Since *Pachylota* has a β-structure and the repeat has an even number of residues, the resulting β-sheets will have one face comprised entirely of alanine residues and the other entirely of glutamine residues. β-sheets in proteins commonly have a right-handed twist, and aggregation with a second β-sheet of similar characteristics is common *in vivo*. In such cases, the apolar faces of two interacting β-sheets (i.e., those composed solely of alanine residues) would be most likely to come together. In so doing, the apolar residues would be shielded from the aqueous milieu and the assembly would gain stability. A repeat of this type would thus have structural implications rather than functional ones.

There are multiple phosphorylation sites in the tail domain of the medium and high molecular weight chain of neurofilaments (NF–H). In the majority of cases, the phosphorylation sites are serine residues within the

repeating triplet motif of exact sequence K–S–P. None of these are contiguously arranged, though pairs of them are found within a 14-residue repeat of the form K–S–P–E–K–A–K–S–P–V–K–E–A–A (http://www.interfil.org/). This is conserved at the 89% level (i.e., a Type C repeat) and occurs nine times in tandem. The K–S–P repeat (Type A), which is primarily a functional one, represents a site of kinase action. This has significant implications *in vivo*. Indeed, phosphorylation/dephosphorylation events are important not only in neurofilaments, but also in intermediate filaments in general. In the former case, the state of phosphorylation seems to relate directly to axonal diameter, conduction velocity, and transport properties.

An example of a longer exact repeat is found in chick scale keratin. The sequence contains a fourfold tandem repeat, each 13 residues long (G–Y–G–G–S–S–L–G–Y–G–G–L–Y; Fig. 1). Interestingly, feather and scale keratin share a common microfibril structure with a 3.4 nm diameter.

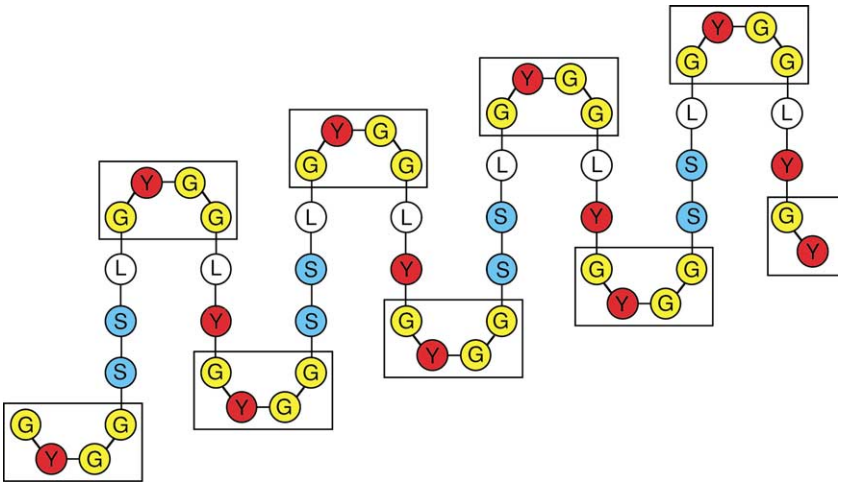


FIG. 1. The fundamental difference between the sequences of scale and feather keratin is the presence of four consecutive repeats of a 13-residue motif in the former but not in the latter. Gregg *et al.* (1984) have proposed that this 52-residue sequence forms an eight-stranded antiparallel β-sheet, likely to have a right-handed twist. A schematic of the conformation illustrates the confinement of glycine residues (in yellow) to the β-turns and the serine residues (in blue) to the β-strands. A tyrosine residue (in red) occurs at each β-turn and in alternate β-strands; these may be involved in forming strong apolar interactions with other tyrosines in similar sheets, thereby providing the scale keratin microfibrils with greater lateral organization, as observed experimentally. Figure redrawn from the original of Gregg *et al.* (1984).

However, the packing of these elements in scale keratin is considerably more regular. This has been associated with the four 13-residue repeats, since the sequences of the two proteins differ only in their inclusion in scale keratin and their absence in feather keratin. According to Gregg *et al.* (1984), the 52-residue region adopts an eight-stranded antiparallel β-sheet structure with the tyrosine residues located at the turns. Tyrosines are often considered to be ''sticky'' residues since they frequently form strong hydrophobic interactions. It can be speculated, therefore, that the more regular packing in scale keratin is specified in large part by interactions between the β-sheets from different molecules. The sequence evidence supports the concept that scale and feather keratin shared a common ancestor. As scales preceded feathers chronologically, the feathers would have evolved from scales through the deletion of the tandem repeat region.

None of the very long sequence repeats found in proteins—perhaps exemplified by the fivefold tandem 534-residue repeat in epiplakin (Fujiwara *et al.*, 2001)—has a structure that has been determined experimentally by either X-ray on nuclear magnetic resonance (NMR) means. Nonetheless, it is possible to infer that sequences over 50 residues in length (let alone those more than 500 residues long) will form a structural domain (or perhaps several domains) to yield a structure that can be likened to a string of beads. This form automatically provides a protein with multiple opportunities to interact similarly with other proteins or with other biological macromolecules. Profilaggrin, a polyprotein precursor to filaggrin, is composed extensively of Type C repeats and can be likened somewhat to epiplakin in this regard (Gan *et al.*, 1990; Rothnagel and Steinert, 1990). Human profilaggrin contains 10–12 tandem repeats, each 312 residues long, and these are joined by short highly conserved linkers. Mouse profilaggrin, however, has even more repeats (18 to > 30) and two distinct types exist (250 and 255 residues in length). It must be emphasized here, though, that profilaggrin is a precursor protein and not a functional one. Furthermore, its repeats are of the Type C, rather than the Type A, variety.

B. Type B Repeats

As noted earlier, the structural motif in Type I collagen α-chains is of the form $(G\text{-}X\text{-}Y)_n$, where X and Y can be almost any amino or imino acid residue and n is 338. This triplet substructure currently represents one of the best examples of a Type B repeat (i.e., a repeat in which one or more residues in a motif is maintained absolutely for structural or functional reasons). Some residues in the X or Y positions of the collagen α-chains,

however, show a preference for one position or the other; this may arise either from stereochemical limitations (see, for example, the preference for leucine, phenylalanine, and glutamic acid in the X positions, and for glutamine, arginine, and lysine in the Y positions: Fietzek and Kühn, 1975; Salem and Traub, 1975) or from posttranslational modifications (see, for example, prolines in the X positions, but hydroxyprolines in the Y positions). Each α-chain has a left-handed conformation closely akin to that seen in polyglycine II (Rich and Crick, 1955). *In vivo*, three such chains aggregate in parallel to form a three-stranded right-handed coiled-coil molecule with 10 residues in three turns (Fig. 2; Fraser *et al.*, 1979, 1983; Ramachandran and Kartha, 1955; Rich and Crick, 1955, 1961).

The structural roles of the glycine residues on the one hand, and the residues in the X and Y positions on the other, are quite distinct. The glycines are a major determinant of the geometry of the triple helix and are internally located close to the axis of the molecule. Their small size makes them the only residue capable of fitting in to the available space without causing a major distortion to the structure of the triple helix.

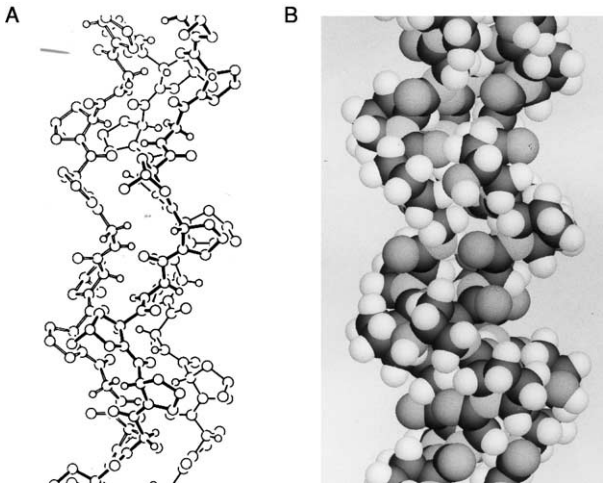


FIG. 2. (A) The molecular structure of collagen contains three left-handed helical α-chains that coil about one another in a right-handed manner. As a consequence of the glycine-based triplet substructure in the collagen sequence, the glycine residues in all three chains lie close to the axis of the molecule, and are the only residues able to fit into the central space available due to their size. The model was produced by Fraser *et al.* (1979) using a linked-atoms least-squares technique refined against quantitative X-ray diffraction data. (B) Space-filling version of the collagen structure illustrated in (A). Figure from Fraser *et al.* (1987) with permission from Academic Press.

In contrast, the X and Y residues lie totally on the surface of the molecule, where they may interact readily with other molecules and hence specify the mode of molecular aggregation. The silk from the gooseberry sawfly *Nematus ribesii* gives a collagen-type X-ray diffraction pattern, which is strongly indicative that a triplet substructure also exists in its amino acid sequence.

C. Type C Repeats

A very large number of Type C repeats have been recorded. Examples of these include: human neurofilament chain (NF–H) with 89% identity among the nine contiguous repeats in a 14-residue consensus sequence K–S–P–E–K–A–K–S–P–V–K–E–E–A (Lee *et al.*, 1988); human ribosome receptor with 87% sequence identity among the 54 contiguous repeats in a 10-residue consensus sequence N–Q–G–K–K–A–E–G–A–Q (Langley *et al.*, 1998); desmoyokin, a keratinocyte plasma membrane-associated protein, with 78% identity among the ten contiguous repeats in a 128-residue consensus sequence (Hashimoto *et al.*, 1993); giardia from median body with 47% identity among the 11 contiguous repeats in a 24-residue consensus sequence (Marshall and Holberton, 1993); and human erythrocyte α-spectrin with 21% identity among 20 contiguous repeats in a 106-residue consensus sequence (Sahr *et al.*, 1990).

A good example of a significant species difference is given by the repeats in the tail domain of nestin (an intermediate filament from neuronal cells). In hamsters, the sequence displays 18 consecutive repeats, with each being 44 residues in length. Human nestin, in contrast, was shown to have 22 residue repeats with only 14 in tandem. The sequences of both repeats, nonetheless, showed a close relationship (Steinert *et al.*, 1999b). The 22-residue repeat has an underlying quasi-repeat of just 11 residues (consensus sequence K/E–E–D/N–Q–E–X–L–R/K–X–L–E).

The seven-residue heptad repeat is one of nature's most frequently used oligomerization motifs. Each repeat has the form $(a\text{–}b\text{–}c\text{–}d\text{–}e\text{–}f\text{–}g)_n$, where positions *a* and *d* are generally filled by apolar residues (occupancy rate is approximately 75% in both positions). The chains adopt an α-helical conformation with about 3.6 residues per turn; since the apolar residues are, on average, 3.5 residues apart, they form a left-handed stripe that winds around the axis of the right-handed α-helix. When two or more α-helices aggregate, the apolar residues become internalized in a knob-hole form of close packing and are shielded from the water. The coiled-coil structures thus formed display a range of pitch lengths. Two-stranded GCN4, for example, has a pitch length of about 20.4 nm (Fig. 3; Harbury *et al.*, 1993; Kühnel *et al.*, 2004), but three- and four-stranded coiled coils

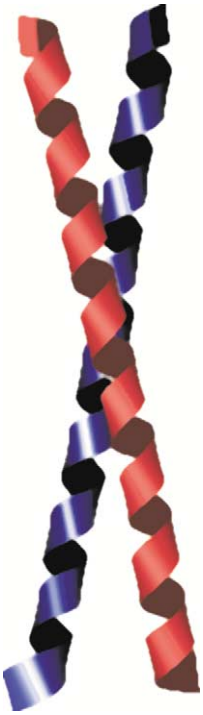


FIG. 3. The structure of the 33-residue region of yeast transcription factor GCN4 is a two-stranded coiled coil, and is viewed here perpendicular to its long axis. The chains each have a heptad substructure and an α-helical conformation. Because GCN4 contains leucine residues in each *d* position, except for the most C-terminal one, the structure is commonly referred to as a leucine zipper (PDB coordinate reference number 2ZTA). The pitch length of the left-handed coiled coil has an average value of about 20.4 nm (Harbury *et al.*, 1993; Kühnel *et al.*, 2004).

generally display somewhat greater values (see methodology of Strelkov and Burkhard, 2002). It is worth reiterating that the quantal repeat (seven residues) does not depend structurally or functionally on the need to have conserved residues in any particular position. The character, however, is strongly maintained through the presence of apolar residues in positions *a* and *d*, and by the frequent occurrence of charged residues (DEKR) in positions *e* and *g*. The latter play an important role in forming interchain ionic interactions that specify both the relative chain stagger and the relative polarity of those chains. The heptad repeat is a very common feature of proteins, most specifically the α-fibrous class. This includes the muscle proteins myosin and tropomyosin, the intermediate

filament proteins, and the plakin superfamily of intermediate filament associated proteins and fibrinogen, among many others.

Four-stranded coiled coils, both with 11-residue repeats, have been recognized in tetrabrachion (residues 19–52; Özbek *et al.*, 2004; Stetefeld *et al.*, 2000) and RH4 protein (Burkhard *et al.*, 2001). These repeats are equivalent to two heptads with a stutter (i.e., there is a three-residue deletion in an otherwise continuous heptad substructure; Brown *et al.*, 1996). On average, the apolar residues are 11/3, or 3.67 residues apart. As this value is slightly greater than the average number of residues per turn in an α-helix, the coiled-coil structure will have a slight right-handed twist; pitch lengths of 128 and 84 nm have been recorded (Kühnel *et al.*, 2004). However, a short length of sequence at the N-terminal end of the tetrabrachion chain (residues 4–18) has a four-residue insert and this, in effect, leads to a 15-residue repeat. This is equivalent to a pair of heptads plus a skip residue. A skip is conformationally equivalent to a pair of stutters. In this instance, the average separation of apolar residues is 15/4, or 3.75 residues. This is considerably larger than the average number of residues per turn in an average α-helix, and the right-handed pitch length for this segment is consequently much shorter (32.2 nm) than those observed for structures with 11-residue repeats. Further support for this conclusion comes from the work of Kühnel *et al.* (2004), who have recently solved the structure of human vasodilator-stimulated phosphoprotein (VASP). This contains a pair of 15-residue repeats with the right-handed pitch length for this structure calculated to be 18.5 nm, a value directly comparable to that seen in two-stranded left-handed coiled coils (Fig. 4).

Another example of a quantal repeat—but with considerable variation in sequence—is seen in the keratin-associated proteins (KAPs). In sheep, these display pentapeptide and decapeptide consensus repeats of the form C–C–Q–P–S/T and C–C–Q/R–P–S/T–C/S/T–C–Q–P/T–S, respectively (Parry *et al.*, 1979). Some of the positions, as indicated by the presence of a consensus sequence, contain residues that occur much more frequently than others, but the absolute conservation of a residue in any position is not observed. The decapeptide consists of a pair of five-residue repeats closely related, but different to that displayed by the pentapeptide. Although the repeats have an undetermined structure, the similarity of the repeat to a sequence in snake neurotoxin suggests that the pentapeptides will adopt a closed loop conformation stabilized by a disulphide bond between cysteine residues four apart (Fig. 5: Fraser *et al.*, 1988; Parry *et al.*, 1979). Relative freedom of rotation about the single bond connecting disulphide-bonded knots would give rise to the concept of a linear array of knots that can fold up to form a variety of tertiary structures. The KAPS display imperfect disulphide stabilization of knots and have interacting

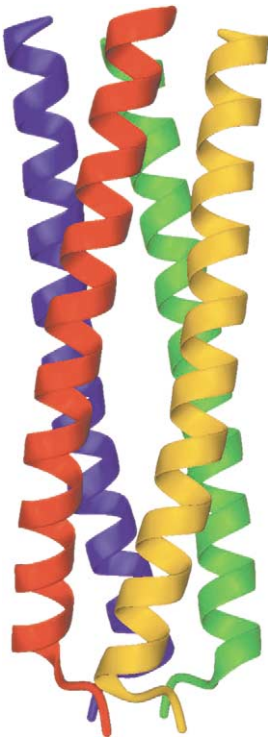


FIG. 4. The structure of human vasodilator-stimulated phosphoprotein (VASP) is that of a parallel chain, four-stranded coiled-coil (Kühnel *et al.*, 2004). The sequence contains a pair of 15-residue repeats, resulting in the formation of a right-handed coiled coil with a pitch length of 18.5 nm, a value directly comparable to that seen in left-handed two-stranded coiled coils. Figure courtesy of Sergei Strelkov.

sites that differ subtly from one another, presumably for functional reasons.

SH3 domains are used extensively by cytoskeletal and signaling proteins to mediate protein-protein interactions, and they do so through a proline-rich motif. This has a consensus sequence P–X–X–P. Another motif—the WW domain—also facilitates protein-protein interactions, and it too is based on proline. Its consensus sequence is P–P–X–Y (a Type I WW repeat as identified in the extracellular matrix receptor β-dystroglycan) and this interacts with a WW domain in dystrophin or utrophin (Ilsley *et al.*, 2002; Winder, 2001). These two motifs are interesting, not only because they are short and proline-rich, but because they are able to impose considerable specificity of interaction on the proteins involved.

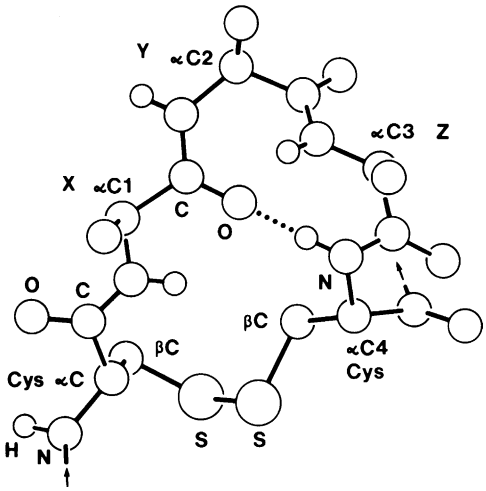


FIG. 5. Predicted conformation of the pentapeptide repeat C–X–Y–Z–C in trichocyte keratin-associated proteins. Glutamine and arginine residues are found commonly in the X position, prolines in the Y position, and serines and threonines in the Z position. The structure is based on the known conformation of a similar repeat in snake neurotoxin. The model shows a disulphide bond-stabilized β-bend with a potential hydrogen bond (dotted). A string of these β-bends, linked by bonds about which there is relatively free rotation, has been proposed as a model for this important family of matrix proteins in trichocyte keratin (Fraser *et al.*, 1988). Figure from Fraser *et al.* (1988) with permission from Elsevier.

D. Type D Repeats

The structure of intermediate filament protein molecules is a tripartite one with globular head and tail domains separated by a rodlike structure of typical length 46 nm. The latter consists of four α-helical heptad-containing coiled-coil segments (1A, 1B, 2A, 2B) connected by linkers L1, L12, and L2 respectively (http://www.interfil.org: Parry and Steinert, 1995). The 1B and 2A + 2B regions both have regular linear distributions of their acidic and basic residues, thus giving rise to alternating acidic and basic residue banding. The observed periods are 9.54 residues (1.42 nm) and 9.84 residues (1.46 nm), respectively. On the basis of maximizing intermolecular ionic interactions, five possible modes of molecular aggregation were proposed by Crewther *et al.* (1983). Three of these modes involved antiparallel alignments and were substantiated experimentally by crosslinking studies (Steinert *et al.*, 1993a,b,c, 1999a). They arose from an alignment in different molecules of: (1) the 1B segments (the A_{11} mode); (2) the 2B segments (the A_{22} mode); and (3) the 1B and 2B segments (the A_{12} mode) (Fig. 6).

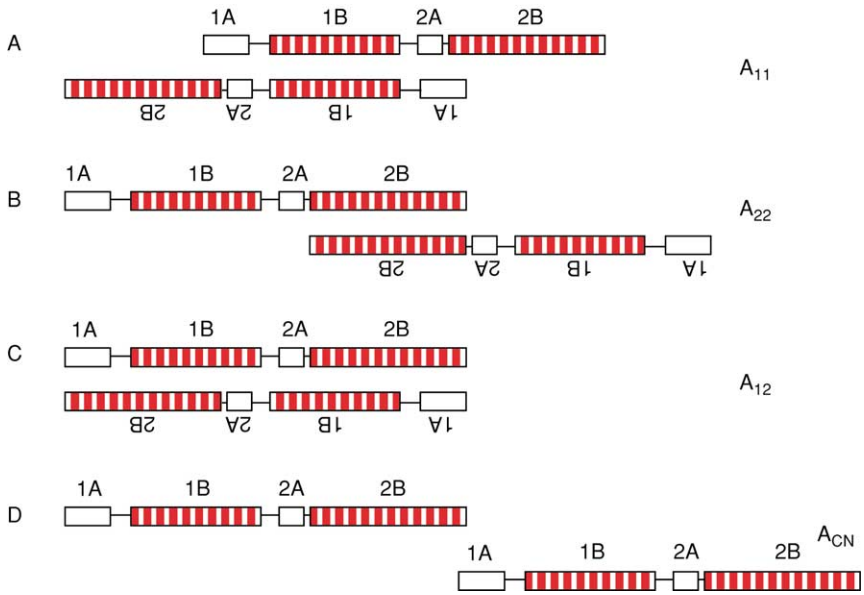


FIG. 6. The longest coiled-coil segments in an intermediate filament chain are known as segments 1B and 2B. Both display a regular distribution of their acidic and their basic residues (period in both charged types is about 9.54 residues in segment 1B, and about 9.84 residues in segment 2B). The bands of opposite charge in these segments are indicated by the red and the white stripes. Alignment of molecules to maximize potential intermolecular ionic interactions can be achieved with either parallel or antiparallel arrangements, though crosslinking data have shown the presence only of the latter. The red bands lie opposite the white bands to indicate one of a family of possible axial staggers that maximize the potential of forming intermolecular ionic interactions. Antiparallel alignment occurs between (A) two 1B segments (A_{11} mode), (B) two 2B segments (A_{22}), and (C) entire molecules (A_{12}). (D) A fourth mode of assembly (A_{CN}) involves a small head-to-tail overlap (typically about seven residues) between similarly directed molecules.

Tropomyosin is a two-stranded, α-helical coiled-coil molecule that aggregates head-to-tail with others to form long filamentous ropes. These lie in each of the two long period grooves of the actin microfilaments where, in vertebrate skeletal muscle, they play an important part in the Ca^{2+}-mediated regulation of actin via troponin (a tropomyosin-associated protein). An important feature of tropomyosin is its 39.2-residue period—that is also quasi-halved (19.6 residues)—in the linear distribution of the acidic residues and, to a lesser extent, the apolar residues (McLachlan and Stewart, 1976; Parry, 1975). The number of residues in tropomyosin (284 residues), and the head-to-tail overlap (nine residues) that allows axial

aggregation to occur, generate an effective repeat length of 275 residues. This is almost exactly equal to seven times the observed sequence repeat ($7 \times 39.2 = 274.4$ residues). In other words, a repeat length of tropomyosin gives rise to a sevenfold copy of a pair of segments with similar sequences (the so-called α- and β-sites). Just as importantly, the 39.2-residue period in tropomyosin ($39.2 \times 0.1485 = 5.82$ nm) corresponds exactly to the separation of the actin monomers in the thin filament. This means that each actin is capable of being regulated by tropomyosin in the same manner (Fig. 7; McLachlan and Stewart, 1976; Parry, 1975; Parry and Squire, 1973; Phillips *et al.*, 1986). This is perhaps a classic example of how proteins with completely different helical symmetries are able to assemble and interact (and function) in a highly specific manner. When the muscle is relaxed, the site on actin that interacts with the myosin head is blocked by the α-site of tropomyosin, and myosin-ATPase activity is not activated. However, when the muscle is activated (through a Ca^{2+} release mechanism), tropomyosin is believed to roll across the surface of the thin filament and thus allow the β-sites to interact with actin. Two possible substrates here, however, have been suggested. One of these allows weak binding of myosin heads to the thin filaments and has low ATPase activity, and the other is a fully active site that allows strong binding of myosin to the thin filaments and gives the greatest ATPase activity (see, for example, Brown *et al.*, 2001). Further details of the regulation of vertebrate skeletal muscle by tropomyosin are given by Squire and Morris (1998).

The α-fibrous proteins myosin (http://motility.york.ac.uk/myosins.shtml) and paramyosin in muscle have structural similarities to one another, with each having a heptad substructure and a resulting long coiled-coil rod domain (about 163 and 122 nm, respectively). *In vivo*, myosin and paramyosin coassemble in such a manner that paramyosin forms the core of the thick filaments and myosin lies on their surfaces (Cohen and Parry, 1998). In order for this to happen, there must be complementary interactions between the molecules; indeed, there are periodic distributions of acidic and basic residues in both myosin and paramyosin. The periods are identical in both proteins for both residue groupings, and in each case these are equivalent to about 9.4 residues. The fundamental period, however, lies close to 28 residues and arises from the beating together of the 28/3 and heptad repeats (Cohen *et al.*, 1987; Kagawa *et al.*, 1989; Mclachlan and Karn, 1983; Parry, 1981). Consequently, appropriate alignment of the acidic and basic residue clumps in the two molecular species (i.e., relative axial staggers of 28 [n + 0.5] residues) allows the formation of many stabilizing intermolecular ionic interactions. This is compatible with the periods seen experimentally in muscle filaments using both X-ray diffraction and electron microscope techniques.

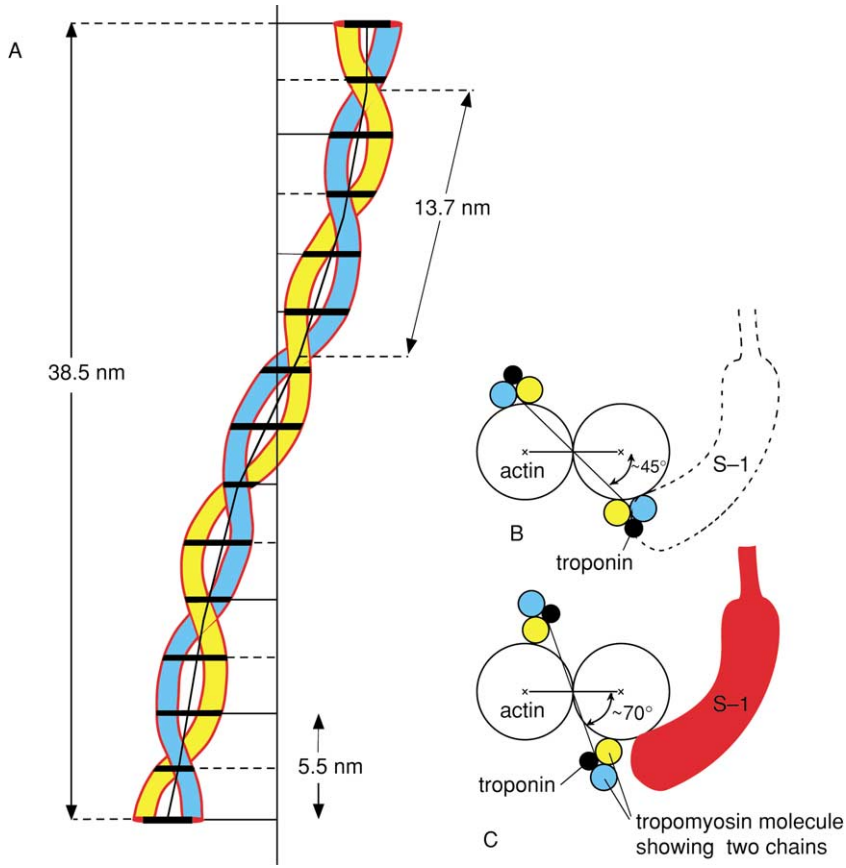


FIG. 7. (A) Tropomyosin contains two right-handed α-helical chains (in blue and yellow) that are parallel and in axial register. These chains coil around one another in a left-handed manner typical of a coiled-coil molecule. In turn, head-to-tail assemblies of parallel tropomosin molecules follow the long period right-handed grooves in the thin filament. Each tropomyosin molecule has 14 zones with an acidic and an apolar part (McLachlan and Stewart, 1976; Parry, 1975). These are represented by horizontal black bars. For a supercoil pitch length of 13.7 nm, tropomyosin will make seven half turns relative to the seven actin monomers with which it makes contact. Alternate zones will have identical azimuthal aspects with regard to each actin, thereby allowing each actin to be regulated by tropomyosin in an identical or quasi-identical manner. If tropomyosin rolls across the actin surface by about 90 degrees, then the second set of zones will lie in roughly equivalent azimuthal positions with respect to the first set prior to the rolling motion. The relaxed state of muscle is thus stabilized by one set of seven tropomyosin-actin interactions and the active state by the second set. (B) Cross-section of the thin filament in relaxed vertebrate skeletal muscle illustrating the positions of the head-to-tail arrays of tropomyosin molecules on the surface of each of the two long-period actin strands. The model was determined using X-ray diffraction

E. Type E Repeats

Type E repeats can have a structural and/or a functional role *in vivo*, and examples of both are given below. Although in many cases the repeats tolerate variation in length, the conformational character is maintained. This enables the motif to aggregate appropriately or function as required. Most of the currently observed Type E repeats lie within the size range of 20–50 residues, but there is no reason to believe that this limitation is mandatory; examples lying outside this range will probably be found.

Calcium-binding proteins are important in intracellular Ca^{2+} signal transduction, in the modulation of Ca^{2+} signals and in Ca^{2+} homeostasis. The calcium binding sites, which are known as EF hands (http://structbio.vanderbilt.edu/chazin/cabp_database), have a consensus sequence (in calmodulin) of the form E–h–X–X–h–h–X–X–h–D–X–D–G–D–G–X–I–D–X–E/D–E–h–X–X–h–h–X–X–h, where residues 1–9 and 22–29 adopt an α-helical conformation and residues 10–21 form a turn that acts as the Ca^{2+} binding site (Fig. 8). The symbol *h* refers to an apolar residue. Small variations from this consensus do occur in some of the numerous Ca^{2+}-binding proteins that have been characterized to date. In particular, some of the EF hands contain a small deletion or insertion with respect to the consensus, thus giving rise to a motif with a length that is a little different from the idealized 29-residue motif noted above. Generally, the EF hands occur in pairs, and these are connected by a short linker. This feature, however, is not well preserved in terms of sequence. Interestingly, the first EF hand in a pair may form one of two types of Ca^{2+} binding loops. In the first, it is 12 residues long, calmodulin-like, and chelates calcium primarily by interactions with side-chain carboxylates. In the second, it is 14 residues long, S100-like and chelates calcium mainly through interactions with backbone carbonyls. In EF hands, there is an appreciable conformational change between the state when Ca^{2+} is bound and the one in which it is not (Fig. 8). In particular, in the former state, a significant hydrophobic surface is exposed, and this allows interaction to occur with a peptide,

---

data (Parry and Squire, 1973). Tropomyosin blocks the normal site of interaction of the head of myosin (S-1) on the thin filament. As S-1 is not able to bind actin, it is shown as dotted. The two parallel in-register α-helical strands of tropomyosin (in yellow and blue) are shown (arbitrarily) to lie in a flat-on position on actin. Troponin is illustrated as a red circle. There is no significance in this figure as to the actual site on tropomyosin at which troponin is shown as binding. (C) As in (B), except that the muscle is now shown in a contracting state, wherein the tropomyosin strands have rolled over the actin surface, possibly by about 90 degrees, to allow the second set of interacting sites on tropomyosin, roughly 2.9 nm distant along the chain, to interact with actin once again in a flat-on position.

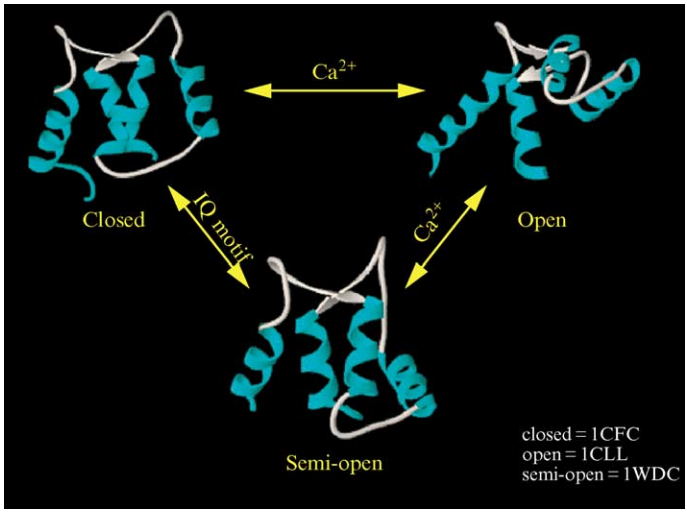


FIG. 8. The EF hand represents a widespread Ca^{2+} binding motif that was first characterized in calmodulin. Generally, EF hands are observed in pairs. Each consists of two short stretches of α-helix connected by a loop. Two possible sets of ligands have been identified through which calcium cations bind to the residues comprising the loop (see text). This figure illustrates the structural forms observed that relate to different Ca^{2+} binding states (PDB: closed, 1CFC; open, 1CLL; semi-open, 1WDC). Figure from http://structbio.vanderbilt.edu.chazin/cabp_database/seq.

often helical (Broderick and Winder, 2002). The Ca^{2+} binding motif is a good example of a functional Type E repeat.

The term *zinc finger* was originally used to describe the multiple zinc-binding motifs with DNA-binding properties in the Xenopus transcription factor IIIA. Nowadays, however, it is used more widely to specify a family of structurally diverse zinc-stabilized compact domains (http://prodata.swmed.edu/zndb). The classic zinc finger, as initially described, was up to 28 residues in length and had a consensus sequence of the form h–X–C–X_{2-5}–C–X_3–h–X_5–h–X_2–H–X_{2-5}–H. Although this motif, like all Type E repeats, displays variability in length, the cysteine and histidine residues remain regularly placed to act as zinc ligands. Two of these ligands were found from residues in the N-terminal β-hairpin structure; the other two were found in the C-terminally-located α-helix. The sidechains of residues that lay close to the N-terminal end of the α-helical strand were shown to be requisites for binding to the major groove of DNA. It is important to emphasize here, however, that at least eight separate zinc finger structures

are now known, each with its own distinct sequence characteristics and conformation, and the various structural forms are described in detail by Krishna *et al.* (2003). The number of zinc fingers per protein varies greatly too, with as few as 2 to as many as 37 (Klug and Schwabe, 1995).

Leucine-rich repeats represent binding motifs found in a wide variety of both plant and mammalian proteins (Kobe and Kajava, 2001). These are involved in a multitude of protein-protein interactions. The sequence of porcine ribonuclease inhibitor, for example, displays a leucine-rich repeat (LRR) of length 27–29 residues that occurs 15 times in tandem (Fig. 9). Likewise, the family of small leucine-rich proteoglycans that includes biglycan, decorin, epiphycan, fibromodulin, keratocan, and lumican

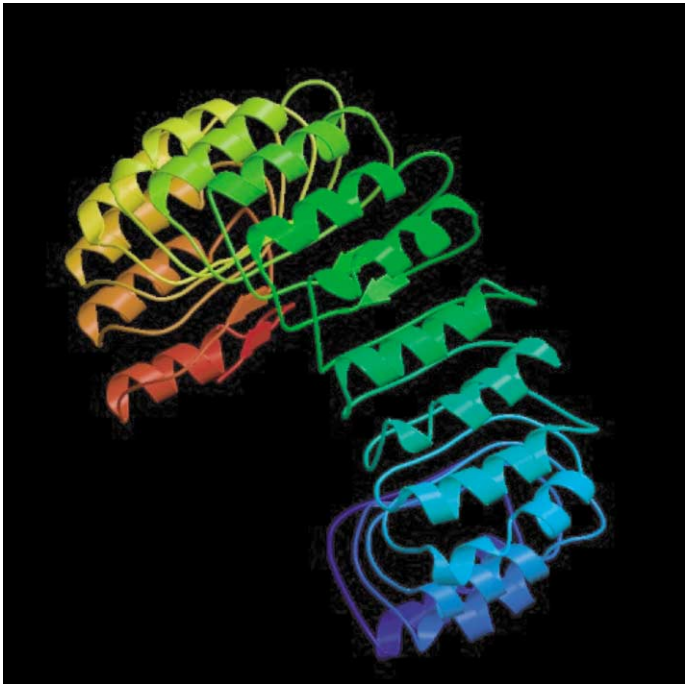


FIG. 9. The conformation adopted by a leucine-rich repeat (LRR) is that of a β-strand followed by an α-helix. In porcine ribonuclease inhibitor, a β-strand (residues 2–8) is connected to an α-helix (residues 14–27) by a connecting loop (residues 9–13). A horseshoe-shaped structure is formed and is exemplified in the crystal structure of ribonuclease inhibitor (PDB 1A4Y: Kobe and Deisenhofer, 1993). This has an inner concave surface formed by curved β-sheets and an outer convex surface formed by α-helices. The leucines and other large apolar residues form the hydrophobic core of the structure.

has a central region in which 8–10 LRRs occur in tandem (Iozzo and Murdoch, 1996). The LRR repeat length can vary somewhat in length and is, for example, a little shorter in decorin (24 residues) than in ribonuclease inhibitor, though it too occurs many times (10 times) in tandem. In other LRR proteins, repeat lengths between 20 and 26 residues have been observed. However, the key feature related to the regular disposition of the leucine residues is maintained. The consensus sequence thus has the general form X–L–X–X–L–X–L–X–X–N/C–X–L–X–X–X–X–X–X–X–L–X–X–X–L–X–X–X, but the first part L–X–X–L–X–L–X–X–N/C–X–L represents the ''signature'' of the LRR. The conformation adopted by the motif, irrespective of its exact length, is that of a β-strand followed by an α-helix. In porcine ribonuclease inhibitor, there are two alternating repeats, but in both cases the β-strand (residues 2–8) is connected to the α-helix (residues 14–27) by a connecting loop (residues 9–13). The horseshoe-shaped structure thus formed and exemplified in the crystal structure of ribonuclease inhibitor (Kobe and Deisenhofer, 1993) has an inner concave surface formed by curved β-sheets and an outer convex surface formed by α-helices. The leucines and other large apolar residues form the hydrophobic core of the structure (Fig. 9).

HEAT-like repeats (http://www.embl-heidelberg.de/~andrade/papers/rep/search.html) have been recognized in two proteasome-binding proteins—PA200 and Ecm29 (Kajava *et al.*, 2004). Much of the sequences of these two proteins can be accounted for in terms of multiple repeats, of which there are about 29 in Pa200 (each about 38–50 residues long) and about 18 in Ecm29 (where the motif is similar in length). These repeats are degenerate and difficult to recognize but, nonetheless, retain their structural identity. In spite of the degeneracy, the character of the repeat in terms of the type of amino acids present in most positions is maintained. The HEAT-like repeats seen here can, in many respects, be considered as an extreme case of ''poor'' definition in what is normally understood by the term *sequence repeat*. Structurally, the motif will adopt an α-helical solenoid conformation, as deduced from knowledge of the structure of 11SREG (a proteasome regulator), a related protein with many HEAT repeats (Whitby *et al.*, 2000).

IV. Summary

Sequence repeats in proteins make use of structural and functional motifs that nature has found to work well *in vivo*. Large numbers of these motifs have now been characterized, and any newly determined amino acid sequence is generally run through computer databases to assess the number and type of repeats present, their possible conformations, and their likely functions *in vivo*. The characteristics of the sequence motifs

differ greatly; variations have been distinguished relating to size, the degree of internal conservation of the sequence, and whether or not the repeats are contiguous or separated from one another.

This review has concentrated on the fibrous proteins. These are, by definition, generally characterized by the presence of regular short-range sequence repeats and (usually) a helical arrangement of those structural motifs, thereby generating a structure with a high axial ratio. In virtually all cases, these elongate molecular structures are interspersed with or terminated by globular domains. Repeats in these regions have also been noted where appropriate.

In the past, there was a tendency to associate structural roles with the rod domains in a fibrous protein, and the functional roles with the globular regions. The recent acquisition of considerable quantities of structural data available from X-ray diffraction and NMR studies has shown that, while this generalization retains some value, the concept should no longer be taken too literally. Many exceptions have now been observed. For example, the observation that globular regions can play a significant structural role *in vivo* is exemplified by the H1 subdomains in the ''globular'' head domains of epidermal keratin chains. These have been shown to stabilize oligomers at the early stages of filament assembly. Likewise, it has been demonstrated that rod domains can play an important functional role, as seen in the regulatory mechanism of vertebrate skeletal muscle. This involves the movement of tropomyosin molecules across the surface of the actin monomers in the thin filaments, thereby providing the mechanism by which actins can all be regulated in a like manner.

Although there are many more sequence repeats in fibrous proteins than have been detailed in this review, it is believed that all of them, together with those noted in some detail in the preceding text, can be classified into one of only five general types. Each has its own structural and functional characteristics, knowledge of which is becoming increasingly useful as the sequences of ever larger proteins become known. Indeed, the bulk of superlong sequences can be accounted for by a linear array of well-characterized motifs, of which (in most cases) the three-dimensional structures are known. Consequently, it is becoming more and more probable that a structure relating to a new protein sequence can be determined in terms of an array of motifs of known conformation. Protein prediction techniques, allied to molecular dynamic calculations, thus have the potential to allow the complete structure to be formulated without the need for X-ray or NMR methods. Although the field is not yet quite at that stage, it may well become a reality within the next decade or so.

References

Broderick, M. J. F., and Winder, S. J. (2002). Towards a complete atomic structure of spectrin family proteins. *J. Struct. Biol.* **137,** 184–193.

Brown, J. H., Cohen, C., and Parry, D. A. D. (1996). Heptad breaks in α-helical coiled coils: Stutters and stammers. *Proteins* **26,** 134–145.

Brown, J. H., Kim, K-H., Jun, G., Greenfield, N. J., Dominguez, R., Volkmann, N., Hitchcock-DeGregori, S. E., and Cohen, C. (2001). Deciphering the design of the tropomyosin molecule. *Proc. Natl. Acad. Sci. USA* **98,** 8496–8501.

Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001). Coiled coils: A highly versatile protein folding motif. *Trends Cell Biol.* **11,** 82–88.

Cohen, C., and Parry, D. A. D. (1986). α-helical coiled-coils–A widespread motif in proteins. *Trends Biochem. Sci.* **11,** 245–248.

Cohen, C., and Parry, D. A. D. (1990). α-helical coiled-coils and bundles: How to design an α-helical protein. *Proteins* **7,** 1–15.

Cohen, C., and Parry, D. A. D. (1998). A conserved C-terminal assembly region in paramyosin and myosin rods. *J. Struct. Biol.* **122,** 180–187.

Cohen, C., Lanar, D. E., and Parry, D. A. D. (1987). Amino acid sequence and structural repeats in schistosome paramyosin match those of myosin. *Biosci. Rep.* **7,** 11–16.

Crewther, W. G., Dowling, L. M., Steinert, P. M., and Parry, D. A. D. (1983). Structure of intermediate filaments. *Int. J. Biol. Macromol.* **5,** 267–274.

Fietzek, P. P., and Kühn, K. (1975). Information contained in the amino acid sequence of the alpha1(I)-chain of collagen and its consequences upon the formation of the triple helix, of fibrils and crosslinks. *Mol. Cell. Biochem.* **8,** 141–157.

Fraser, R. D. B., MacRae, T.P., and Suzuki, E. (1979). Chain conformation in the collagen molecule. *J. Mol. Biol.* **129,** 463–481.

Fraser, R. D. B., MacRae, T. P., Miller, A., and Suzuki, E. (1983). Molecular conformation and packing in collagen fibrils. *J. Mol. Biol.* **167,** 497–521.

Fraser, R. D. B., MacRae, T. P., Chew, M. W. K., and Squire, J. M. (1987). Collagen and Elastin. *In* ''Fibrous Protein Structure'' (J. M. Squire and P. J. Vibert, Eds.), pp. 173–191. Academic Press, London.

Fraser, R. D. B., MacRae, T. P., Sparrow, L. G., and Parry, D. A. D. (1988). Disulphide bonding in α-keratin. *Int. J. Biol. Macromol.* **10,** 106–112.

Fujiwara, S., Takeo, N., Otani, Y., Parry, D. A. D., Kunimatsu, M., Lu, R., Sasaki, M., Matsuo, N., Khaleduzzaman, M., and Yoshioka, H. (2001). Epiplakin, a novel member of the plakin family originally identified as a 450-kDa human epidermal autoantigen: Structure and tissue localization. *J. Biol. Chem.* **276,** 13340–13347.

Gan, S.-Q., McBride, O. W., Idler, W. W., Markova, N. G., and Steinert, P. M. (1990). Organization, structure, and polymorphisms of the human profilaggrin gene. *Biochemistry* **29,** 9432–9440.

Gregg, K., Wilton, S. D., Parry, D. A. D., and Rogers, G. E. (1984). A comparison of genomic coding sequences for feather and scale keratins: Structural and evolutionary implications. *EMBO J.* **3,** 175–178.

Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two-, three- and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262,** 1401–1407.

Hashimoto, T., Amagai, M., Parry, D. A. D., Dixon, T. W., Tsukita, S., Miki, K., Sakai, K., Inokuchi, Y., Kudoh, J., Shimizu, N., and Nishikawa, T. (1993). Desmoyokin, a

680 kDa keratinocyte plasma membrane-associated protein, is homologous to the protein encoded by human gene AHNAK. *J. Cell Sci.* **105,** 275–286.

Hulmes, D. J. S. (1992). The collagen superfamily – Diverse structures and assemblies. *Essays Biochem.* **27,** 49–67.

Hulo, N., Sigrist, C. J. A., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P., and Bairoch, A. (2004). Recent improvements to the PROSITE Database. *Nucl. Acids Res.* **32,** D134–D137.

Ilsley, J. L., Sudol, M., and Winder, S. J. (2002). The WW domain: Linking cell signaling to the membrane cytoskeleton. *Cell. Signal.* **14,** 183–189.

Iozzo, R. V., and Murdoch, A. D. (1996). Proteoglycans of the extracellular environment: Clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* **10,** 598–614.

Kajava, A. V., Gorbea, C., Ortega, J., Rechsteiner, M., and Steven, A. C. (2004). New HEAT-like repeat motifs in proteins regulating proteasome structure and function. *Trends Biochem. Sci.* (in press).

Kagawa, H., McLachlan, A. D., Brenner, S., Karn, J., and Gengyo-Ando, K. (1989). Paramyosin gene (unc-15) of *Caenorhabditis elegans*: Molecular cloning, nucleotide sequence and models for thick filament structure. *J. Mol. Biol.* **207,** 311–333.

Klug, A., and Schwabe, J. W. R. (1995). Zinc fingers. *FASEB J.* **9,** 597–604.

Kobe, B., and Deisenhofer, J. (1993). Crystal structure of the porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature (London)* **366,** 751–756.

Kobe, B., and Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **11,** 725–732.

Krishna, S. S., Majumdar, I., and Grishin, N. V. (2003). Structural classification of zinc fingers. *Nucl. Acids Res.* **31,** 532–550.

Kühnel, K., Jarchau, T., Wolf, E., Schlichting, I., Walter, U., Wittinghofer, A., and Strelkov, S. V. (2004). The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proc. Natl. Acad. Sci. USA* **101,** 17027–17032.

Langley, R., Leung, E., Morris, C., Berg, R., McDonald, M., Weaver, A., Parry, D. A. D., Ni, J., Su, J., Gentz, R., Spurr, N., and Krissansen, G. W. (1998). Identification of multiple forms of 180-kDa ribosome receptor in human cells. *DNA Cell Biol.* **17,** 449–460.

Lee, V. M.-Y., Otvos, L., Carden, M. J., Hollosi, M., Dietzschold, B., and Lazzarini, R. A. (1988). Identification of the major multiphosphorylation site in mammalian neurofilaments. *Proc. Natl. Acad. Sci. USA* **85,** 1998–2002.

Lucas, F., and Rudall, K. M. (1968). Variety in composition and structure of silk fibroins: Some new types of silk from the hymenoptera. *In* ''Symposium on Fibrous Proteins, Australia 1967'' (W. G. Crewther, Ed.), pp. 45–55. Butterworths, Sydney.

Lupas, A. (1996). Coiled coils: New structures and new functions. *Trends Biochem. Sci.* **21,** 375–382.

Marshall, J., and Holberton, D. V. (1993). Sequence and structure of a new coiled-coil protein from a microtubule bundle in *Giardia. J. Mol. Biol.* **231,** 521–530.

McLachlan, A. D. (1978). Coiled coil formation and sequence regularities in the helical regions of α-keratin. *J. Mol. Biol.* **124,** 297–304.

McLachlan, A. D., and Karn, J. (1983). Periodic features in the amino acid sequence of nematode myosin rod. *J. Mol. Biol.* **164,** 605–626.

McLachlan, A. D., and Stewart, M. (1976). The 14-fold periodicity in α-tropomyosin and the interaction with actin. *J. Mol. Biol.* **103,** 271–298.

McLachlan, A. D., and Stewart, M. (1982). Periodic charge distribution in the intermediate filament proteins desmin and vimentin. *J. Mol. Biol.* **162,** 693–698.

Özbek, S., Müller, J. F., Figgemeier, E., and Stetefeld, J. (2004). Insights into thermostability at atomic level: The crystal structure of the right-handed tetrabrachion coiled coil of *Staphylothermus marinus* at 1.4 A resolution. *J. Mol. Biol.* (in press).

Parry, D. A. D. (1975). Analysis of the primary sequence of α-tropomyosin from rabbit skeletal muscle. *J. Mol. Biol.* **98,** 519–535.

Parry, D. A. D. (1979). Determination of structural information from the amino acid sequences of fibrous proteins. *In* ''Fibrous Proteins: Scientific, Industrial, and Medical Aspects'' (D. A. D. Parry and L. K. Creamer, Eds.), Vol. 1, pp. 393–427. Academic Press, London.

Parry, D. A. D. (1981). Structure of rabbit skeletal myosin: Analysis of the amino acid sequences of two fragments from the rod region. *J. Mol. Biol.* **153,** 459–464.

Parry, D. A. D., and Squire, J. M. (1973). The structural role of tropomyosin in muscle regulation: Analysis of the X-ray diffraction patterns from relaxed and contracting muscles. *J. Mol. Biol.* **37,** 33–55.

Parry, D. A. D., and Steinert, P. M. (1995). Intermediate Filament Structure. Springer-Verlag, Heidelberg.

Parry, D. A. D., Crewther, W. G., Fraser, R. D. B., and MacRae, T. P. (1977). Structure of α-keratin: Structural implication of the amino acid sequences of the Type I and Type II chain segments. *J. Mol. Biol.* **113,** 449–454.

Parry, D. A. D., Fraser, R. D. B., and MacRae, T. P. (1979). Repeating patterns of amino acid residues in the sequences of some high-sulphur proteins from α-keratin. *Int. J. Biol. Macromol.* **1,** 17–22.

Phillips, G. N., Fillers, J. P., and Cohen, C. (1986). Tropomyosin crystal structure and muscle regulation. *J. Mol. Biol.* **192,** 111–131.

Ramachandran, G. N., and Kartha, G. (1955). Structure of collagen. *Nature (London)* **176,** 593–595.

Rich, A., and Crick, F. H. C. (1955). The structure of collagen. *Nature (London)* **176,** 915–916.

Rich, A., and Crick, F. H. C. (1961). The molecular structure of collagen. *J. Mol. Biol.* **3,** 483–506.

Rothnagel, J. A., and Steinert, P. M. (1990). The structure of the gene for mouse filaggrin and a comparison of the repeating units. *J. Biol. Chem.* **265,** 1862–1865.

Sahr, K., Laurila, P., Kotula, L., Scarpa, E., Coupal, E., Leto, T., Linnenbach, A., Winkelmann, J., Speicher, D., Marchesi, V., Curtis, P., and Forget, B. (1990). The complete cDNA and polypeptide sequences of human erythroid α-spectrin. *J. Biol. Chem.* **265,** 4434–4443.

Salem, G., and Traub, W. (1975). Conformational implications of amino acid sequence regularities in collagen. *FEBS Lett.* **51,** 94–99.

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., and Bucher, P. (2002). PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief Bioinform.* **3,** 265–274.

Squire, J. M., and Morris, E. P. (1998). A new look at thin filament regulation in vertebrate skeletal muscle. *FASEB J.* **12,** 761–771.

Steinert, P. M., Marekov, L. N., Fraser, R. D. B., and Parry, D. A. D. (1993a). Keratin intermediate filament structure: Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J. Mol. Biol.* **230,** 436–452.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1993b). Conservation of the structure of keratin intermediate filaments: Molecular mechanism by which different keratin molecules integrate into pre-existing keratin intermediate filaments during differentiation. *Biochemistry* **32,** 10046–10056.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1993c). Diversity of intermediate filament structure: Evidence that the alignment of coiled-coil molecules in vimentin is different from that in keratin intermediate filaments. *J. Biol. Chem.* **268,** 24916–24925.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1999a). Molecular parameters of Type IV α-internexin and Type IV-Type III α-internexin-vimentin copolymer intermediate filaments. *J. Biol. Chem.* **274,** 1657–1666.

Steinert, P. M., Chou, Y.-H., Prahlad, V., Parry, D. A. D., Marekov, L. N., Wu, K. C., Jang, S.-I., and Goldman, R. D. (1999b). A high molecular weight intermediate filament-associated protein in BHK-21 cells is nestin, a Type VI intermediate filament protein: Limited co-assembly *in vitro* to form heteropolymers with Type III vimentin and Type IV α-Internexin. *J. Biol. Chem.* **274,** 9881–9890.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R. A. (2000). Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–776.

Strelkov, S. V., and Burkhard, P. (2002). Analysis of α-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J. Struct. Biol.* **137,** 54–64.

Whitby, F. G., Masters, E. I., Kramer, L., Knowlton, J. R., Yao, Y., Wang, C. C., and Hill, C. P. (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature (London)* **408,** 115–120.

Winder, S. J. (2001). The complexities of dystroglycan. *Trends Biochem. Sci.* **26,** 118–124.

THE STRUCTURE OF α-HELICAL COILED COILS

By ANDREI N. LUPAS AND MARKUS GRUBER

Max-Planck-Institute for Developmental Biology, D-72076 Tübingen, Germany

I. Historical Introduction 37
II. Structural Parameters 40
A. The Standard Model 40
B. Prediction and Analysis Programs 45
C. Discontinuities 49
D. Coiled Coils that Deviate Globally from the Standard Model 53
III. Structural Determinants of Folding and Stability 56
A. Number of Helices 56
B. Orientation and Oligomer Specificity 57
C. Folding and Stability 58
IV. Structural Diversity 60
A. Fibers and Zippers 60
B. Tubes, Sheets, Spirals, Funnels, and Rings 61
C. What Exactly is a Coiled Coil? 67
V. Function Follows Structure 70
References 72

Abstract

α-Helical coiled coils are versatile protein domains, supporting a wide range of biological functions. Their fold is probably better understood than that of any other protein; indeed, uniquely among folds, their structure can be computed from a set of parametric equations. Here, we review the principles of coiled-coil structure, the determinants of their folding and stability, and the diversity of structural forms they assume.

I. Historical Introduction

The first investigations into the structure of coiled coils were made by William Astbury at the University of Leeds in the 1930s. Astbury had worked with Sir William Bragg at the Royal Institution in London in the mid 1920s and, at Bragg's request, had obtained X-ray diffraction patterns for wool and silk. After moving to Leeds in 1928, he began applying X-ray diffraction systematically to a large spectrum of natural fibers and related materials. His main interest was textile fibers, particularly wool in its native

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70003-0

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

(unstretched) and denatured (stretched) forms, since Leeds is located at the heart of Britain's textile industry. He also studied many protein fibers from other sources (such as porcupine quills, horse hair, horn, and tendons) and discovered the existence of three main spectral forms: an α-form exemplified by unstretched wool, a β-form exemplified by stretched wool, and a γ-form corresponding to tendon. All three forms showed a high degree of fiber regularity and became the targets of numerous modeling studies. The α-form was the most common of the three spectral forms; Astbury named the class of proteins showing this diffraction pattern *k-m-e-f*, for keratin, myosin, epidermin, and fibrinogen. It showed strong meridional arcs at 5.15 Å (indicating the repeating unit of the structure) and a group of equatorial reflexions at around 10 Å and again at 27 Å.

From 1948 on, Linus Pauling at the California Institute of Technology and Lawrence Bragg's group at the Cavendish Laboratory in Cambridge became engaged in a race for the correct structural model of the α-form. Pauling won in 1950 with the announcement of two hydrogen-bonded helical structures (Pauling and Corey, 1950; Pauling *et al.*, 1951), one of which was the α-helix (the other structure, a helix with 5.1 residues per turn, has never been observed in nature). The α-helix had a rise of 1.5 Å per residue and 3.6 residues per turn, yielding a periodicity of 5.4 Å, not 5.15 Å. Pauling had ignored the meridional arcs at 5.15 Å because of the diffraction pattern of a synthetic fiber, poly-γ-methyl-L-glutamate, which was clearly of the α-type but lacked the meridional arcs, instead showing relexions at 5.4 Å away from the meridian. The α-helix was soon confirmed by Perutz (1951), who discovered the 1.5 Å diffraction spot required by this structure; it had hitherto been missed because of the size of the photographic plates and the angle between fiber and X-ray beam used for measurements. The prominent meridional arc of k-m-e-f proteins, however, remained unexplained.

It seemed clear that, while the α-helix was the dominant structural element in proteins with an α-type diffraction pattern (hence the 'α' in α-helix), it needed to be further modified to obtain the fiber structure. This insight led both Crick and Pauling to consider superhelical distortions arising from the packing of neighboring α-helices, reports of which they submitted independently of each other to *Nature* at the end of 1952 (Crick, 1952; Pauling and Corey, 1953). In his description of the α-helix (Pauling *et al.*, 1951), Pauling had dismissed the need for the number of residues per turn to be an integer or a ratio of small integers, but had envisaged the possibility that in a crystalline arrangement, the regular packing of the helices could favor deformations into configurations with

a rational number of residues per turn, such as 11/3 (11 residues over 3 turns), 15/4, and 18/5. The idea that packing interactions would lead to deformations of the α-helix appears to anticipate the key insight needed for the coiled coil, but at this time Pauling was clearly still thinking in terms of nonnative, crystalline interactions and not of native interactions in a fiber. Conspicuously absent from his list of periodicities is 7/2, the signature periodicity of the coiled coil. Two years later, he took up this idea again, proposing that the superhelical distortion arising from α-helices twisting around each other could account for the meridional arc in the diffraction spectrum of k-m-e-f proteins (Pauling and Corey, 1953). He envisaged a range of possible sequence periodicities (4/1, 7/2, 15/4), leading to supercoils with senses of twist both the same and opposite to those of the constituent helices. He also envisaged various stoichiometries for the helical bundles (which he called *compound helices*), including one with six helices coiling around a straight seventh one. Unlike his earlier paper on the α-helix, this paper did not offer quantitative parameters for the model structures. Also, side-chain packing played no role in his considerations. Instead, supercoiling resulted from the exact repetition of short sequences, which caused periodic fluctuations in backbone hydrogen-bond lengths.

In contrast, Crick's communication in *Nature* and two subsequent papers in *Acta Crystallographica* offered a detailed and fully parameterized model for a sequence periodicity of 7/2 (Crick, 1952, 1953a,b). He showed that, if α-helices were to twist around each other at an angle of about 20 degrees, their side chains would interlock systematically along the core of the structure, repeating the same interactions every 7 residues (or two turns of the α-helix). The close packing interactions of hydrophobic residues in the core would provide the energy required to distort the helices, a remarkable insight at a time when the biophysics of protein folding were unknown and even the exact sequence of a protein still remained to be determined. Crick called the bundle of supercoiled helices a *coiled coil* and referred to the regular packing of side chains as ''knobs'' into ''holes.'' The coiled coils would contain two or three parallel α-helices, supercoiled in the opposite sense of the helices. (Note that since Pauling's initial α-helix model was left-handed, Crick's coiled coil was right-handed, although he also considered bundles of right-handed α-helices and of right- and left-handed combinations, stating that the question of helical handedness ''must still be regarded as open.''). Crick's model and the diversity of coiled-coil structures determined in the 50 years since, some of which surprisingly match Pauling's conjectures, form the basis of this Chapter.

II. Structural Parameters

A. *The Standard Model*

Coiled coils are bundles of α-helices that are wound into superhelical structures (Fig. 1). Most commonly, they consist of two, three, or four helices, running in the same (parallel) or in opposite (antiparallel) directions, but structures with five and more helices have been determined. They are usually oligomers either of the same (homo) or of different chains (hetero), but on occasion consist of consecutive helices from the same polypeptide chain, which in that case almost always have an antiparallel orientation.

As proposed by Crick (1952, 1953a,b), the coiled coil's hallmark is the distinctive packing of amino acid side chains in the core of the bundle, called *knobs-into-holes*, in which a residue from one helix (knob) packs into a space surrounded by four side chains of the facing helix (hole). This geometry contrasts with the more irregular packing of helices in globular proteins, often referred to as *ridges-into-grooves* (Chothia *et al.*, 1977), in which a residue packs above or beneath the equivalent residue from the facing helix. (See the structure of spectrin in Fig. 2b for an example of a protein in which two helices interact via knobs-into-holes, while the third is packed via ridges-into-grooves.) Although knobs-into-holes packing is often also referred to as *in register*, this is not strictly true because the amino acid side chains do not point straight away from the helix but are angled towards the amino-terminus. In antiparallel coiled coils, an optimal interaction of side chains is therefore obtained when they point towards each other and their C_α carbons are out of register.

The regular meshing of side chains in knobs-into-holes packing requires that they occupy periodically equivalent positions along the bundle interface. This is not possible with undistorted α-helices, which have approximately 3.6 residues per turn (Chothia *et al.*, 1981); the position of side chains on their surface drifts continuously. By giving the right-handed α-helices a left-handed twist, coiled coils effectively reduce the number of residues per turn to 3.5 with respect to the supercoil axis, and thus allow the position of side chains to repeat after two turns (or seven residues). The residues engaged in knobs-into-holes interactions are usually hydrophobic, whereas the outer residues are hydrophilic; as already surmised by Crick (1953b), the sequence of coiled coils therefore shows a heptad repeat in the chemical nature of side chains. Schematically, the seven structural positions are labeled *a-g*, with *a* and *d* denoting the hydrophobic residues (Fig. 2c).

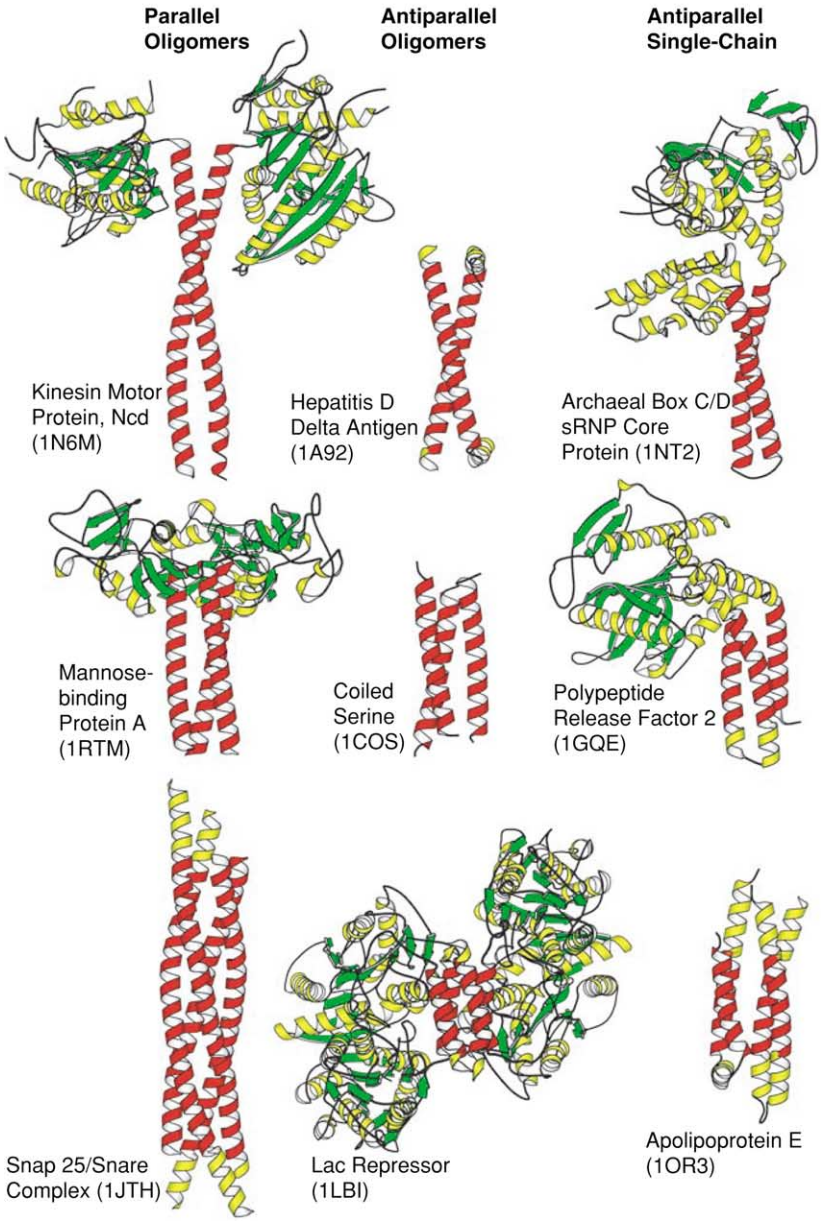


FIG. 1. Gallery of canonical, heptad-based coiled-coil proteins. Coiled-coil segments are shown in red, other α-helices in yellow, and β-strands in green.

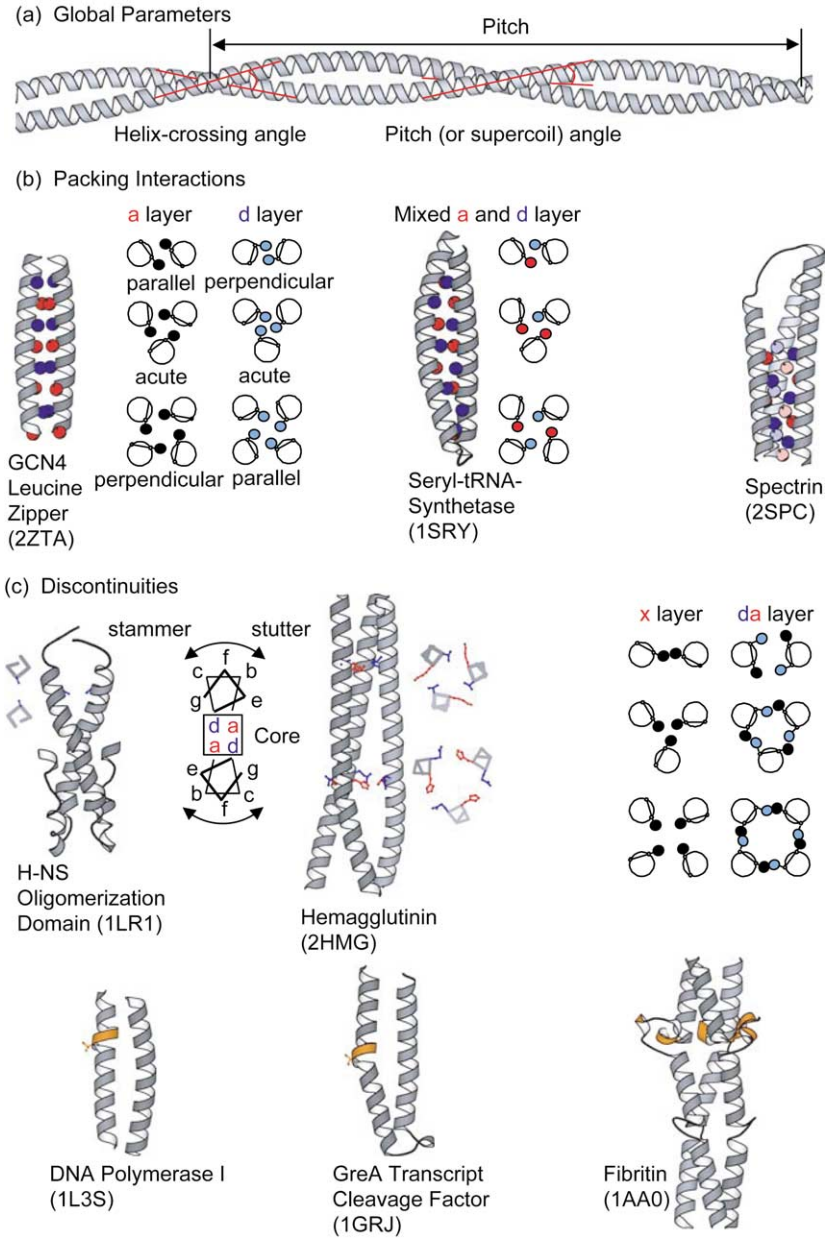


FIG. 2. Parameters of coiled-coil structure. The schematic cross sections are viewed from the amino-terminal end of the structures.

The orientation of hydrophobic side-chains in the core of coiled coils is different for positions *a* and *d* (Fig. 2b), as first described by Harbury *et al.* (1993). In a two-stranded, parallel coiled coil, the C_α–C_β bond of a residue in position *a* is parallel to the peptide bond facing it in the opposing helix; in position *d*, it is perpendicular. Thus, the distance between the C_β carbons of core residues in position *a* (*a* layers) is large, favoring β-branched side-chains for tight packing (Ile, Val, Thr), whereas it is small for core residues in position *d* (*d* layers), favoring residues unbranched at C_β (Leu, Ala). This situation is exactly reversed in parallel, four-stranded coiled coils. In three-stranded coiled coils, the angle is intermediate, leading to a packing orientation called *acute*. In antiparallel coiled coils, core packing layers consist of residues in both *a* and *d*, thus showing mixed geometries (Fig. 2b).

The idealized structure of the coiled coil has been parameterized by Crick (Figs. 2a and 3). The distance required for the superhelix to complete a full turn is called the pitch (*P*), and the angle of a helix relative to the superhelical axis is called the pitch angle (α) [also called

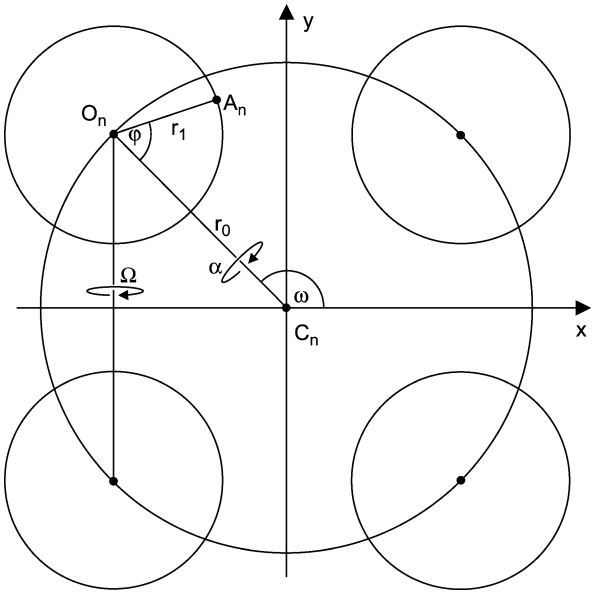


FIG. 3. Schematic representation of a tetrameric coiled coil, showing the main parameters. O_n marks the center of one α-helix, A_n the C_α position of a constituent residue, and C_n the superhelix axis; r_1 is the α-helix radius, r_0 the superhelix radius, α the pitch angle, Ω the pairwise helix-crossing angle, φ the positional orientation angle of a residue or phase of the helix, and ω is the phase of the supercoil.

superhelix crossing angle, χ (Harbury *et al.*, 1994) or tilt angle (Offer and Sessions, 1995)]. The angle between two neighboring helices is the pairwise helix-crossing angle Ω; in two-stranded coiled coils, Ω equals 2α. The vector connecting the center of a helix to the superhelical axis gives the superhelix radius (r_0) and that connecting the center of a helix to the C_α carbons of its constituent residues gives the α-helix radius (r_1). The angle between the r_1 vectors for two consecutive residues is the phase shift of the α-helix ($\Delta\varphi$) and the angle between two consecutive r_0 vectors is the phase shift of the supercoil ($\Delta\omega$); the angle between the r_0 and r_1 vectors for the same residue is the positional orientation angle (φ; Harbury *et al.*, 1993) or Crick angle (α; Strelkov and Burkhard, 2002), which gives the location of a residue relative to the supercoil axis (the choice of α is somewhat unfortunate since α already denotes the pitch angle in Crick's parameterization). The Cartesian coordinates for a coiled coil according to Crick are:

$$\begin{aligned} x &= r_0 \cdot \cos\omega - r_1 \cdot \cos\omega \cdot \cos\phi + r_1 \cdot \cos\alpha \cdot \sin\omega \cdot \sin\phi \\ y &= r_0 \cdot \sin\omega - r_1 \cdot \sin\omega \cdot \cos\phi - r_1 \cdot \cos\alpha \cdot \cos\omega \cdot \sin\phi \\ z &= P(\omega/2\pi) + r_1 \cdot \sin\alpha \cdot \sin\phi \end{aligned} \tag{1}$$

In transforming the coordinates x, y, z of an ideal α-helix (e.g., polyalanine; Arnott and Dover, 1967), these equations can be represented for chain i of an n-stranded coiled coil as:

$$\begin{aligned} x_i &= r_0 \cdot \cos(2\pi z/P + \omega_i) + x \cdot \cos(2\pi z/P + \omega_i) - y \cdot \cos\alpha \cdot \sin(2\pi z/P + \omega_i) \\ y_i &= r_0 \cdot \sin(2\pi z/P + \omega_i) + x \cdot \sin(2\pi z/P + \omega_i) + y \cdot \cos\alpha \cdot \cos(2\pi z/P + \omega_i) \\ z_i &= z - y \cdot \sin\alpha \end{aligned} \tag{2}$$

where $\omega_i = 2\pi(i-1)/n$ and the pitch angle α is:

$$\alpha = \arctan(2\pi r_0/P) \tag{3}$$

Using the formula of Fraser and McRae (1973), the pitch P can be calculated from the supercoil radius r_0, the axial rise per amino acid h (1.495 Å for polyalanine; Arnott and Dover, 1967), and the twist differential Δt:

$$P = (2\pi/\Delta t) \cdot [h^2 - (r_0 \Delta t)^2]^{1/2} \tag{4}$$

where Δt is derived from the number of residues per turn of an undistorted α-helix ($a = 3.62$ for polyalanine; Arnott and Dover, 1967) and the periodicity of hydrophobic residues ($p = 3.5$ in a canonical coiled coil) as:

$$\Delta t = 2\pi \cdot (1/a - 1/p) \tag{5}$$

In the structure of coiled coils, the values for pitch and crossing angle follow directly from the degree of distortion necessary to reach a periodically recurring position for the core residues; Phillips (1992), Seo and

Cohen (1993), and Strelkov and Burkhard (2002) have described methods for computing the local pitch from coordinates (see also the addendum by Zhang and Hermans (1993) and errata in the same issue). Undistorted helices have about 3.64 residues per turn (Chothia *et al.*, 1981) and a rise of 1.50 Å per residues; for two-stranded coiled coils, these values yield a pitch of approximately 140 Å and a helix crossing angle of 22 degrees (Seo and Cohen, 1993). An analysis of various crystal structures (Table I) shows that individual coiled coils range around these values with comparatively minor variations; however, because the pitch is very sensitive to these parameters, even sequences that are nearly identical may yield structures with substantially different pitch. In the two-, three- and four-stranded GCN4 leucine zipper variants, the undistorted helices have 3.65, 3.64, and 3.60 residues per turn, respectively, resulting in pitches of 135 Å, 163 Å, and 188 Å. In addition, because of variations in the sequence and discontinuities in the heptad pattern, the local pitch may vary substantially along the length of a coiled coil, occasionally even leading to reversals in handedness (Seo and Cohen, 1993; Strelkov and Burkhard, 2002).

B. Prediction and Analysis Programs

The strong heptad periodicity of coiled coils and the clear and simple parameterization of their structures have made possible a large number of computational approaches to their analysis. The earliest sequence–based approach was Fourier transform, used initially to detect higher–order periodicities superimposed on the basic heptad pattern (see e.g., Parry, 1975; McLachlan and Karn, 1983; McLachlan and Stewart, 1976), but subsequently also for detecting deviations from the heptad pattern itself (Hoiczyk *et al.*, 2000; Marshall and Holberton, 1993; Peters *et al.*, 1996). In this context, McLachlan developed a Fortran implementation, SEQFFTX, based on a derivation presented by McLachlan and Stewart (1976). The current availability of this program is not known to us. We have implemented a C program for the same purpose (FTWin) which, unlike SEQFFTX, uses a user-determined scanning window; the program can be downloaded from our web site (http://protevo.eb.tuebingen.mpg.de/download).

A second widely used sequence-based approach relies on matrices of residue frequencies, pioneered by Parry (1982). He showed that the residue distribution at the seven heptad positions of the putative coiled-coil segments of myosin, tropomyosin, α-keratin, and hemaglutinin are asymmetric, and proposed a method by which the residue frequencies could be used to predict whether a sequence of unknown structure would form a coiled coil. This approach was implemented with modifications in

TABLE I
Values for the Main Structural Parameters in Coiled Coils with Various Periodicities, Determined Using the Program TWISTER

Periodicity		Number of strands	cc radius r_0 (Å)	cc pitch (Å)	Residue phase (°)	Residues per turn	Rise per residue (Å)
7	GCN4-p1 (2ZTA)	2	4.85	134.5	a: 22.91 d: −32.18	3.62	1.51
	GCN4-pII (1GCM)	3	6.76	163.4	a: 17.72 d: −31.13	3.60	1.53
	GCN4-pLI (1GCL)	4	7.14	187.9	a: 20.65 d: −28.69	3.58	1.52
11	Tetrabrachion (1FE6: 19–48)	4	7.59	11589.7		3.67	1.49
15	YadA (1P9H)	3	6.54	165.0		3.70	1.51
	Tetrabrachion (1FE6: 5–18)	4	7.38	197.5		3.66	1.51
18	Hemaglutinin pH4 (1HTM)	3	6.75	379.4		3.63	1.50
25	Phosphoprotein of Sendai Virus (1EZJ)	4	7.82	342.9		3.63	1.51

the Coils program (http://www.ch.embnet.org/software/COILS_form.html); the main modifications concern the substitution of residue preferences for frequencies, introduction of a scanning window, and a procedure to scale scores against reference databases in order to obtain probabilities (Lupas, 1996a; Lupas *et al.*, 1991). A variant of this approach, using pairwise residue correlations, was developed by Berger and colleagues with the program PairCoil (http://paircoil.lcs.mit.edu/cgi-bin/paircoil; Berger *et al.*, 1995); LearnCoil, a further variant that can be trained iteratively on a set of target proteins (Berger and Singh, 1997), is available on the Internet for two protein sets, histidine kinases (http://learncoil.lcs.mit. edu/cgi-bin/learncoil; Singh *et al.*, 1998) and viral membrane fusion proteins (http://learncoil-vmf.lcs.mit.edu/cgi-bin/vmf; Singh *et al.*, 1999). We have found, however, that using pairwise residue correlations leads to an overfitting of the method to the training set; correspondingly, PairCoil and its derivatives are substantially less sensitive than Coils in detecting proteins unlike those in the training set (i.e., multihelical, antiparallel, short, or irregular structures; Lupas, 1996b). The most recent approach is based on Hidden Markov models (MARCOIL; http://www.isrec.isb-sib.ch/BCF/Delorenzi/Marcoil/index.html; Delorenzi and Speed, 2002) and operates without a scanning window, thus removing one of the limitations of Coils and PairCoil. All current methods are designed to detect unbroken heptad repeats.

Matrices of residue frequencies have also been used to discriminate between two- and three-stranded coiled coils. Again, the first matrices showing a difference in residue preferences were compiled by Parry and coworkers (Conway and Parry, 1990, 1991); they observed a clear drop in the proportion of charged residues in the hydrophobic core of three-stranded coiled coils, a loss of bias for basic residues in *a* and acidic residues in *d*, as well as a more general loss of positional preferences for most amino acids. These observations can be interpreted in terms of the increased size and decreased solvation of the hydrophobic core in three-stranded structures, as well as in terms of the more uniform (acute) geometry of core residues (Fig. 2b). Based on such matrices, Woolfson and Alber (1995) developed the program Scorer, and Wolf *et al.* (1997) developed the program MultiCoil (http://multicoil.lcs.mit.edu/cgi-bin/multicoil). Both appear to operate with a fair measure of success; however, we are not aware of any benchmark of the two programs on current structure databases. A basic problem in their operation results from the fact that they use suboptimal detection routines to identify potential coiled-coil regions (Scorer uses Coiler, a program that must recognize at least four heptads by their hydrophobic pattern and a minimum of 65% Leu, Ile, Val, and Ala in positions *a* and *d* to provide an output; MultiCoil uses PairCoil). In addition, the

programs cannot tell when the coiled coil under study is neither two- nor three-stranded. One basic question that has not been analyzed to date is the influence of the intracellular versus extracellular environment on the residue frequencies in the two types of structure. Almost as a rule, long oligomeric coiled coils are dimeric if they are found inside the cell and trimeric if they are outside. Thus, some of the success in distinguishing the two structures may come from overall differences in residue distribution, rather than from differential preferences in the heptad positions.

The main structure-based programs used for the analysis of coiled coils are Socket (http://www.biols.susx.ac.uk/Biochem/Woolfson/html/coiled-coils/socket/; Walshaw and Woolfson, 2001) and Twister (Strelkov and Burkhard, 2002). In addition, Seo and Cohen (1993) have made their program for the computation of pitch and handedness available (contact http://ccohen@brandeis.edu). Socket is designed to detect knobs-into-holes packing in helical bundles, as the most direct way to evaluate the compatibility of a structure with the standard model. The program operates by representing side chains by their center of mass; they are classified as knobs if they contact four or more side-chain centers within a specified distance cutoff (set to 7.0 Å by default). Incidentally, the program assigns an orientation, a register, and the number of constituent helices for each detected coiled coil. Most of the examples for coiled-coil structures shown in this Chapter were chosen from a scan of the Protein Data Bank (PDB) using Socket. The second program, Twister, is designed to list the local structural parameters of coiled-coil structures, based on Crick's parameterization. Twister uses four consecutive C_α carbons to determine the position of the center for each helix, which then determine the location of the supercoil axis. Once the axes are traced, all other parameters follow (Strelkov and Burkhard, 2002). Twister is very well suited to track local fluctuations in coiled-coil structures, but is limited by the requirement that the helices be in a parallel orientation. The parameters in Table I were computed using Twister.

The modeling of coiled-coil structures from sequence-based predictions can be performed with BeammotifCC (Offer *et al.*, 2002), which uses generalized parametric equations in order to produce coordinates for the main chains of coiled coils, even in cases where their sequences differ locally or globally from the heptad pattern. Test models of various coiled coils from PDB (including GrpE, influenza hemagglutinin at low pH, and tetrabrachion) matched the experimentally determined coordinates with a deviation of less than 1.1 Å in backbone atoms, illustrating the accuracy with which the parameterization can replicate the structure of real proteins. Coiled-coil structures have also been modeled successfully using established, template-based modeling techniques (Nilges and Brunger,

1991; O'Donoghue and Nilges, 1997), but these are neither more accurate, nor do they have the computational simplicity and elegance of the parametric approach.

C. *Discontinuities*

Although coiled coils are generally very regular structures, only few are entirely without discontinuities (Fig. 2c). These can be represented as insertions of one or more residues into the heptad pattern. Insertions of one residue are called *skips*, three-residue insertions are *stammers*, and four-residue insertions are *stutters*. Other insertions may also occur, but are not frequent enough to have been named. Insertions can be accommodated structurally by perturbations in backbone continuity or in side-chain packing. Basically, insertions of three or four residues are close enough to a full turn of the helix (3.6 residues) to allow for their accommodation within the helical structure; however, they distort the knobs-into-holes interactions in the core. By inserting more than 3.5 residues, stutters raise the local sequence periodicity, leading to an unwinding of the left-handed supercoil; stammers have the opposite effect. Both have the same effect, however, on the packing geometry in the core: they alter the relative position of residues by changing the degree of supercoiling from the ideal value given by the standard model. Schematically, this can be illustrated as a rotation of the helices in a helical wheel diagram (Fig. 2c): stutters shift residues in position *a* towards the center of the core, resulting in a geometry called an *x* layer, while they shift residues in *d* out of the core and the following residues towards position *a*, resulting in a ring of interacting residues around a central cavity (a *da* layer). The situation for stammers is analogous, except that residues in position *d* yield the *x* layers, while the *da* layers are formed by residues in positions *g* and *a*. In both cases, the knobs-into-holes packing is transformed locally into a knobs-to-knobs interaction. If the insertion occurs close to a core position, that layer is the main one being distorted; however, if it occurs between core positions, two consecutive layers may be distorted (leading to the observation of consecutive *x* and *da* layers). By virtue of pointing towards the supercoil axis, *x* layers are constrained in the size of the side chains they can accommodate, particularly in the case of two-stranded coiled coils where it has been surmised that only alanine or glycine might fit (Lupas, 1995; see also the alanine-containing *x* layer in the histone-like protein H-NS in Fig. 2c). However, a number of crystal structures illustrate the strategies by which two-stranded coiled coils compensate for this steric constraint and allow for the presence of other side chains (Fig. 4). The CbnR transcription factor accommodates its two *x* layers (Asp67 and Ser78)

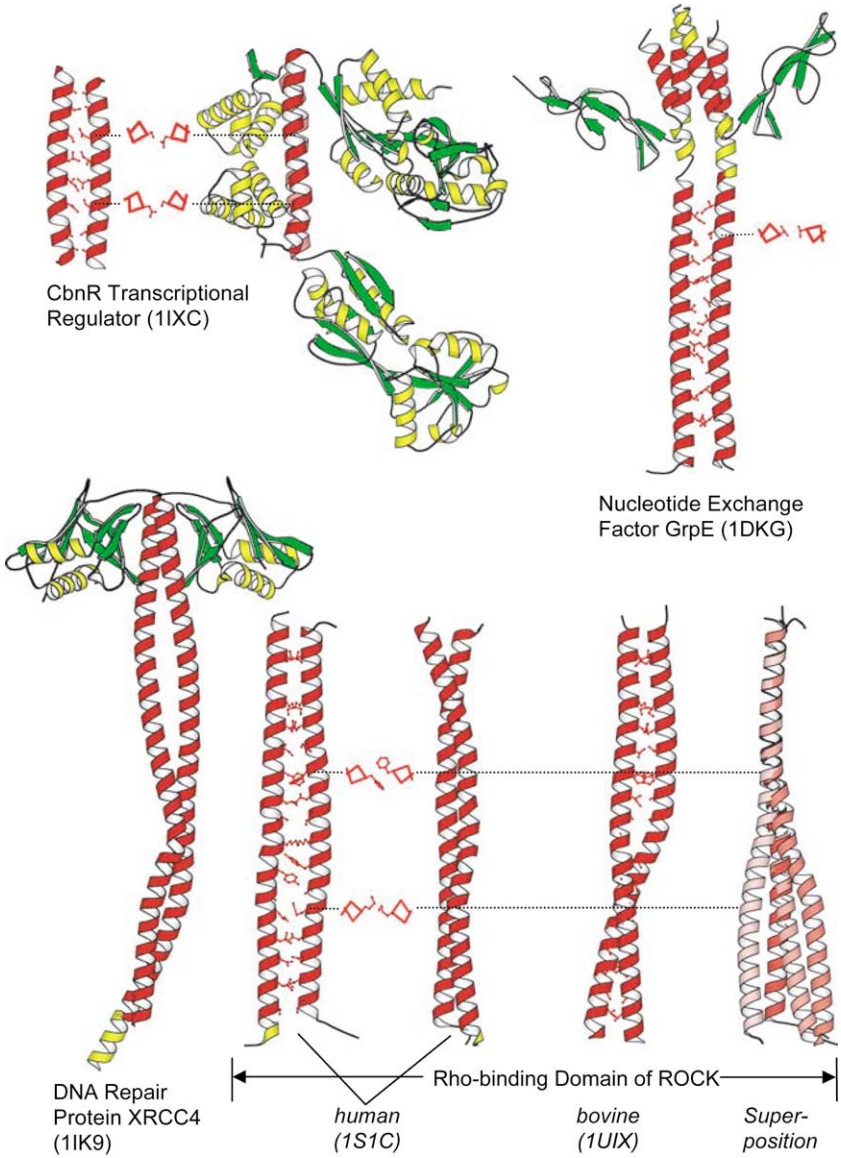


FIG. 4. Constrained knobs-to-knobs layers in two-stranded coiled coils. The figure illustrates the strategies that coiled coils employ in order to accommodate *x* layers in their structure: antiparallel orientation (CbnR), register shifts (GrpE), and symmetry breaks (XRCC4 and ROCK). The constrained core layers are shown in cross-section next to their place in the coiled coils.

by forming an antiparallel structure; it thus combines a larger and smaller side-chain in both *x* layers and orients them past each other (side chains are angled towards the N-terminus and thus point in opposite directions). In other proteins, all of them homooligomeric, *x* layers are accommodated by symmetry breaks. In the chaperone cofactor GrpE, the *x* layers cause the two helices to move out of register, even transforming locally the knobs-into-holes packing into ridges-into-grooves. In the DNA repair protein XRCC4, the *x* layer causes a change of direction of one helix relative to the other. This effect is even more pronounced in the Rho-binding domain of ROCK, in which the helices of two homologs (human and bovine) take four different paths in order to accommodate two *x* layers, going so far as to switch from left-handed to right-handed supercoiling and back (this is also observable, to a lesser extent, in CbnR). The location of the *x* layers is readily seen from the changes in direction of the helices.

Insertions other than stutters or stammers cannot be accommodated without altering the local helical structure. This occurs by preserving as far as possible the hydrogen-bonding pattern and knobs-into-holes packing of the helices and by looping out the extra residues. This is seen most dramatically in fibritin from bacteriophage T4 (Fig. 2c), where insertions of 4 and 12 residues are extruded with minimal disruptions in the continuity of the helix. The extrusion of the four-residue insertion is surprising, given the arguments made above, but may have an energetic reason: both insertions are flanked by glycines, which are sterically required for the extrusion. Conversely, glycines are fairly rare in canonical coiled coils and do not favor an α-helical structure. We think that replacement of the glycines with other residues may cause the incorporation of the four-residue insert as a stutter. More frequent are insertions of just one residue (skips), which may also be looped out to form a π-turn as in the thumb of DNA polymerase I, or accommodated through a kink in the helix as in the GreA transcript cleavage factor (Fig. 2c). Skip residues tend to occur in antiparallel, single-chain coiled coils, which are often more irregular; we are not aware of any skip residues in parallel, oligomeric coiled coils of known structure. Many skip residues have been predicted in coiled coils from sequence analysis; in general, however, most seem to be structurally better represented by two successive stutters (Brown *et al.*, 1996; Lupas *et al.*, 1995; McLachlan and Karn, 1983; Seo and Cohen, 1993), which is not surprising since in a heptad reference frame, the insertion of twice four residues is equivalent to the insertion of a single residue. Indeed, one may argue that pairs of stutters arise most efficiently by the insertion of one residue, whose effects are then delocalized along the helix.

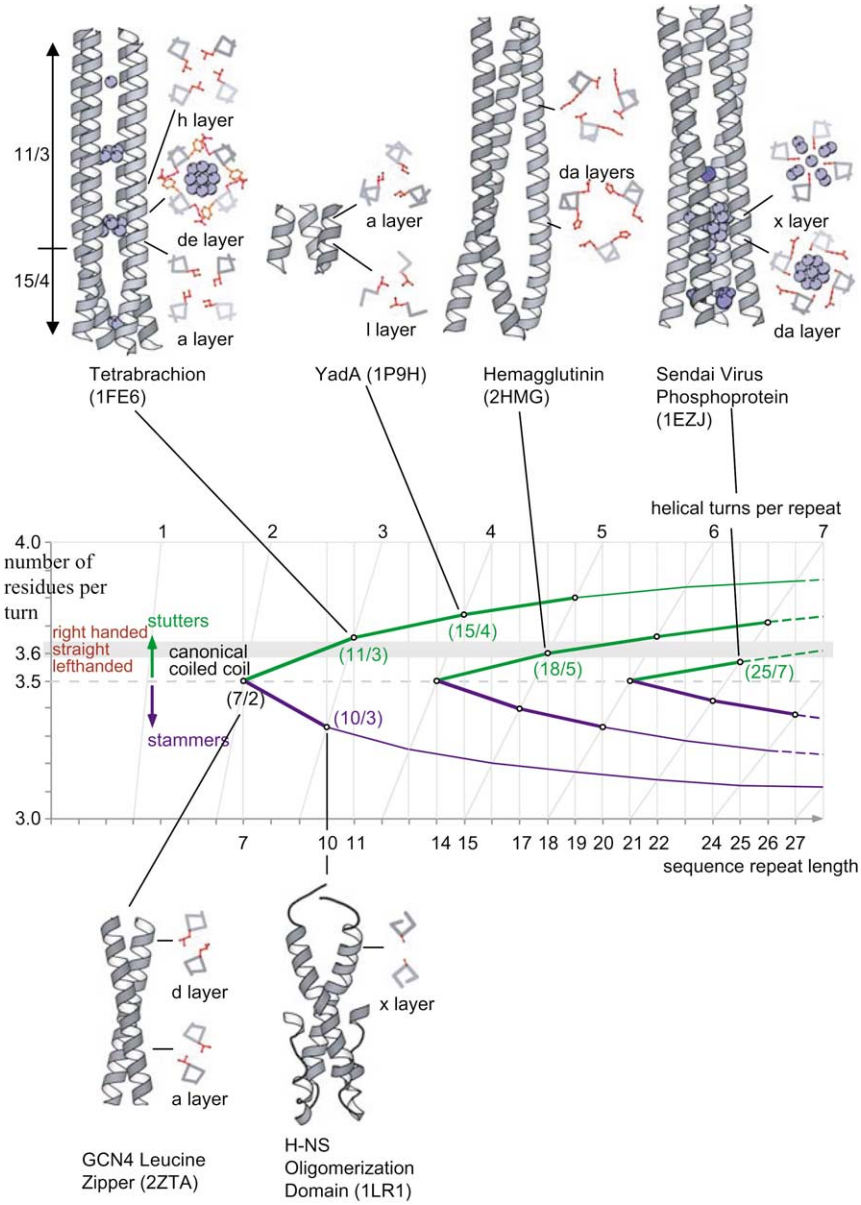


FIG. 5. Schematic diagram of coiled-coil periodicities in relation to the numbers of residues per turn, the sequence repeat length, and the number of helical turns per repeat. Transitions caused by the insertion of stutters are marked in green, and transitions caused by stammers in blue. Periodicities for which we found examples in

D. *Coiled Coils that Deviate Globally from the Standard Model*

Some coiled coils seem to contain discontinuities in a periodically recurring pattern. For example, Holberton and colleagues detected periodicities of 29, 25, and 24 residues in the cytoskeletal proteins of *Giardia* (Holberton *et al.*, 1988; Marshall and Holberton, 1993, 1995) and interpreted these as four heptads plus a skip residue (29), three heptads plus a stutter (25), and three heptads plus a stammer (24). In these proteins, the discontinuities seemed to be spaced sufficiently far apart to suggest structural models with local changes in the pitch and a meandering path of the helices in the coiled coil. Such local changes in pitch angle have been described for a number of structures (Seo and Cohen, 1993; Strelkov and Burkhard, 2002), on occasion even leading to transitions from left-handed to right-handed supercoils, as shown in Fig. 4. The crystal structure of a protein with a 25-residue periodicity, the stalk of Sendai virus phosphoprotein (Fig. 5), shows however that the nonheptad periodicity may be accommodated by more global adjustments of the pitch, corresponding to an overall periodicity of 3.57 (25 residues over 7 turns).

The possibility of such global departures from the heptad model was first suggested by the discovery of proteins with 11-residue periodicities (hendecads), such as the Lea proteins of plants (Dure III, 1993) and the stalk of the archaeal surface protein tetrabrachion (Peters *et al.*, 1996). In these proteins, the overall hydrophobic periodicity of 3.67 (11/3) exceeded the periodicity of an undistorted α-helix (3.63), suggesting a right-handed twist for the supercoil. A detailed analysis of the tetrabrachion stalk by biophysical and bioinformatic techniques revealed the presence of a continuous homotetrameric coiled coil of 70 nm length and extreme stability, which consisted of two segments separated by a single proline residue (Peters *et al.*, 1995, 1996). The N-terminal segment showed a heptad periodicity with two closely spaced stutters (left-handed supercoil), while the C-terminal segment had a clear hendecad periodicity with one additional stutter and three stammers (right-handed supercoil). In modeling the tetrabrachion stalk, the hendecads were treated as a heptad plus a stutter; based on the effects of stutters on packing interactions described above, this meant the addition of a *da* or *x* layer to each heptad, depending on where the stutter was accommodated structurally

---

the protein sequence database are connected by bold lines. Examples for the main periodicities from the structure database are shown above and beneath the diagram with cross sections illustrating core packing layer geometries. The region in which the constituent helices of a coiled coil are expected to be essentially straight is marked with a gray bar; periodicities above this bar cause a right-handed and periodicities beneath a left-handed supercoil.

(using the heptad notation, the eleven positions of a hendecad, *a-k*, are equivalent to *abcdabcdefg* or to *abcdefgxefg*). An analysis of the average hydrophobicity of hendecad positions yielded high values for positions *a* and *h*, and intermediate values for positions *d* and *e*, pointing to a core formed by *a, d, e,* and *h*. As confirmed by the crystal structure of a tetrabrachion fragment (Stetefeld *et al.*, 2000; Fig. 5), positions *a* and *h* of the hendecad correspond structurally to positions *a* and *d* of the heptad, while the hendecad *d* and *e* positions form a knobs-to-knobs packing layer equivalent to a *da* layer. The same core layers are seen in the short coiled coils of HdeA (1BG8), which consists of a single hendecad, and of the estrogen receptor ligand-binding domain (1PCG); 1PCG contains three hendecads, but diverges at the ends, so that only the central hendecad (and, to a lesser extent, the N-terminal hendecad) form regular core interactions. The second possibility, to give *h* an *x*-layer geometry and use three positions (*a, d,* and *h*) as the hydrophobic core, was realized in a designed protein (RH4) with the nonnatural amino acid alloisoleucine in position *d*. We are aware of one natural protein employing this packing mode: transcription factor CbnR (1IXC; Fig. 4), which accommodates the *x* (*h*) layers of its two hendecads by assuming an antiparallel orientation, as discussed in the previous section.

The right-handed part of the tetrabrachion stalk contains a stutter within the hendecad frame, which introduces a 15/4 element (or pentadecad) and locally increases the right-handed supercoil by raising the periodicity to 3.75 (Peters *et al.*, 1996; Stetefeld *et al.*, 2000; Fig. 5). In fact, given helices with 3.63 residues per turn, pentadecads are as strongly supercoiled to the right as heptads are to the left. The major adhesin of Yersinia, YadA, contains a stalk formed almost entirely of pentadecads (Hoiczyk *et al.*, 2000); the crystal structure of the YadA head domain contains part of the first pentadecad of the stalk, revealing a structure similar to the stutter region of tetrabrachion, albeit three- rather than four-stranded (1P9H; Fig. 5). Two-stranded coiled coils formed by pentadecads have also been determined: the tetramerization region of Mnt repressor is an unusual dimer of dimers, formed by two antiparallel coiled coils, each two pentadecads long (1QEY; Nooren *et al.*, 1999). Homing endonucleases (e.g., 1G9Z) contain a short parallel, homodimeric coiled coil of just 12 residues, whose supercoil angle suggests an underlying pentadecad periodicity. For such short coiled coils, side-chain packing does not allow us to distinguish between hendecads and pentadecads, but the comparison between HdeA and homing endonucleases, which are of the same length and favor the same core residues (Gly in *d*, aromatic in *e*, Gly or Ala in *h*) reveals a clearly larger supercoil angle in endonucleases. As in the CbnR repressor discussed earlier, the constrained nature of the *x*-like knobs-to-knobs layers found in Mnt repressor seems to favor an antiparallel

orientation of the helices, as the main alternative to symmetry breaks. In homing endonucleases, the short length of the coiled coil and a preference for Gly and Ala in the knobs-to-knobs core layers obviate these constraints.

Hendecads and pentadecads are part of a larger picture (Gruber and Lupas, 2003). Database scans for putative coiled-coil sequences yield a fair number of instances in which the primary periodicity is not 7/2; the more common ones are 11/3, 15/4, 17/5, and 20/6. A plot of these periodicities shows that they are all related to the heptad repeat by regular insertions of three or four residues (i.e., stammers and stutters; Fig. 5). Indeed, the structure database holds examples for a surprising number of nonheptad periodicities, including 10/3, 11/3, 15/4, 18/5, and 25/7. In all cases, the periodicity can be decomposed into elements of three and four residues, whose hydrophobic pattern and residue composition is fundamentally compatible with the basic coiled-coil structure (Hicks *et al.*, 1997), and Crick's parameterization can be adapted easily to this generalization (Harbury *et al.*, 1998; Offer *et al.*, 2002; Peters *et al.*, 1996). Alternating elements (3–4 and 4–3) yield knobs-into-holes packing, whereas consecutive elements of the same kind (3–3 and 4–4) yield knobs-to-knobs packing. Since four residues are more than a helical turn and three residues are less, and since the α-helix is right-handed, it follows that 4–4 patterns bias the structure towards a right-handed supercoil and 3–3 patterns towards a left-handed one. The graph in Fig. 5 shows how the progressive insertion of 4–4 patterns raises the average number of residues per turn, until a coiled coil crosses the rubicon at around 3.63 and switches to a right-handed supercoil; the progressive insertion of 3–3 patterns has the opposite effect.

It is unclear how many successive elements of the same kind can be used to obtain a coiled coil, but we have seen no evidence of coiled coils with periodicities of 4/1 or 3/1, suggesting that there are limits to the degree of distortion to which α-helices can be subjected. Indeed, most of the nonheptad periodicities we observed bring the coiled coils closer to the unsupercoiled state, rather than increasing the degree of supercoiling; thus, the 7/2 and 15/4 periodicities seem to bracket the main range in which coiled coils are found. [Interestingly, 7/2 and 15/4 were the main periodicities proposed by Pauling and Corey (1953), whose opposite handedness they also recognized.] One may envisage two energetic effects: the penalty incurred by distorting the α-helices into a supercoil and the one incurred by deviating from knobs-into-holes packing. In two-stranded coiled coils, which represent the large majority of all structures, the packing penalty outweighs the supercoil penalty, so that most show periodicities close to 7/2 and thus to the lower end of the bracket. Three- and particularly four-stranded coiled coils do not incur the same penalties for knobs-to-knobs packing, due to the greater amount of space available in

their core layers. Correspondingly, they show much more frequently periodicities deviating from the canonical heptad repeat. (As an aside, they are also harder to detect and analyze, since most computational tools have problems with sequences and structures that deviate from the standard model).

III. Structural Determinants of Folding and Stability

A. *Number of Helices*

The oligomeric state of a coiled coil is determined by packing interactions, and thus depends primarily on the nature of residues in positions *a*, *d*, *e*, and *g*. The geometry of side chains in positions *a* and *d* differs systematically between two-, three-, and four-stranded coiled coils, as discussed in Section II.A; correspondingly, they also have different side chain preferences. In two-stranded coiled coils, the core packing geometry favors β-branched residues in *a* and unbranched or γ-branched residues in *d*; in four-stranded coiled coils, the situation is exactly reversed. Three-stranded coiled coils, in contrast, have a more uniform geometry in both *a* and *d* layers (called *acute*), which makes them the least selective. In a series of experiments with the GCN4 leucine zipper, Harbury *et al.* (1993, 1994) showed that the structure could be switched between two-, three-, and four-stranded states by directed changes in the core residues. Isoleucine in *a* and leucine in *d* produced dimers and the reverse produced tetramers; all other combinations of isoleucine, leucine, and valine produced trimers or, in some cases, mixtures of dimers and trimers. A subsequently obtained retro-GCN4 structure, which represents the inverted sequence of the GCN4 zipper and thus reverses the positions of the core residues in the heptad pattern, unsurprisingly yielded a tetramer (Mittl *et al.*, 2000).

The GCN4 sequence contains an additional oligomerization determinant: a single polar residue (Asn) in position *a*. As Harbury *et al.* (1993) found, mutation of this residue to valine caused a loss of structural specificity and led to mixtures of dimers and trimers. Further mutations at this position to glutamine, lysine, norleucine, and aminobutyric acid (Gonzales *et al.*, 1996a,b), as well as the related investigation of two lysines in position *a* of the Fos leucine zipper (Campell *et al.*, 2002) yielded inconsistent results, suggesting that polar residues may promote oligomerization specificity but that their effect is sensitive to the surrounding structural environment. In general, increased hydrophobicity at the helical interface disfavored dimer formation. These findings are consistent with the differences in residue frequencies for two- and three-stranded

coiled coils observed by Conway and Parry (1990, 1991), namely a strong decrease in the proportion of charged residues in the hydrophobic core of three-stranded coiled coils, as well as a more general loss of positional preferences for most amino acids. One may speculate that the three-stranded state represents a default setting for the coiled coil, and that dimers and tetramers require specific structural determinants for their formation.

Much less is known about the role of residues in positions *e* and *g* in specifying oligomer states. Mutation of residues in position *e* of a leucine zipper peptide to alanine resulted in the formation of tetramers (Krylov *et al.*, 1994), as did the disruption of an interchain ionic interaction in the trimeric coiled coil of cartilage matrix protein (Beck *et al.*, 1997). However, it is currently difficult to do more than speculate on the possibility that increased hydrophobicity and/or decreased side-chain size could promote trimer and tetramer formation: short side chains by being less able to shield the hydrophobic core in a dimeric structure, and hydrophobic side chains by favoring extended packing interactions. Finally, we think that knobs-to-knobs core layers also favor the formation of higher oligomers because of their constrained nature in two-stranded coiled coils, but direct evidence for this conjecture is lacking.

B. Orientation and Oligomer Specificity

The orientation and preference for homo- or heterooligomeric association of helices in a coiled coil is determined primarily by ionic and polar interactions between residues in positions *e* and *g*, and also by interactions between these residues and polar residues in *a* and *d* [see, for example, the Fos-Jun heterodimer (Glover and Harrison, 1995) in comparison with the Jun homodimer (Junius *et al.*, 1996)]. Most interactions occur between residues in *g* of one heptad and *e* of the next heptad (interactions between positions *e* and *g* of the same heptad being generally prevented by the shape of the core). In four-stranded coiled coils, positions *b* and *c* may also contribute, as here the core is much broader. The same interactions are also critical in determining whether a coiled-coil helix will form homooligomers or will preferentially heterooligomerize with a different helix. These effects have been studied in great detail with designed peptides and are discussed in this volume by Woolfson. There are, conversely, rather few results with natural sequences, probably because their low degree of redundancy makes the interpretation of outcomes difficult.

Comparative analyses of the number of favorable interactions in various arrangements have been made to distinguish the register and oligomer specificity of coiled coils, such as tropomyosin (McLachlan and Stewart,

1975), keratin (Parry *et al.*, 1977) and laminin (Beck *et al.*, 1993), and have also been used to predict the heterodimerization partners for leucine zippers (Vinson *et al.*, 1993). Extending these efforts, a quantitative method based on pairwise residue interactions for positions *a, d, e,* and *g* has been developed by Fong *et al.* (2004) and is available on the web (http://compbio.cs.princeton.edu/bzip/). The inclusion of positions *a* and *d* is not only justified by the already mentioned possibility of *a*-*g′* and *d*-*e′* interactions, but also by some (albeit few) results suggesting an influence of the hydrophobic interactions: in the GCN4 leucine zipper, mutation of the core asparagine (Asn16) to alanine results in an antiparallel trimer, which replaces the expected Ala–Ala core layer with a Leu–Leu–Ala and an Ala–Ala–Leu layer, presumably to reduce the effects of cavity formation in the core (Holton and Alber, 2004). Monera *et al.* (1996) have described a similar alanine mutation, which switches the orientation of helices in a designed tetramer in order to form mixed Leu–Ala core layers. Core packing is probably also responsible for the antiparallel nature of the synthetic peptide coiled Ser (Lovejoy *et al.*, 1993), whose single tryptophan residue is too bulky to fit three times into the same core layer.

An interesting light is thrown on the relationship between designed and natural sequences by Arndt *et al.* (2000), who used libraries based on the sequences of Jun and Fos to randomize positions *e* and *g* to Gln, Glu, Arg, and Lys and position *a* to Asn or Val, prior to selecting for heterodimers. They observed an enrichment of favorable interactions in the selected sequences, but even the best heterodimer they identified, WinZip-A1B1, still contained two repulsive *e*-*g* interactions, mirroring much more closely the situation observed in natural sequences than the perfect matches engineered into designed peptides. A further study attempting to optimize WinZip-A1B1 by genetic selection and, alternatively, by rational design found that the best genetically obtained pair retained the repulsive *e*–*g* interactions and that it was more effective than the designed pair in mediating heterodimerization *in vivo* (Arndt *et al.*, 2002). The authors concluded that the effects of predicted charge pairs depend on sequence context, and complementary charges at the edge positions rationalize only a fraction of the sequences that form stable, specific coiled coils.

C. Folding and Stability

Fragments of natural coiled coils, even when they are of considerable size, rarely fold [see, for example, Trybus *et al.* (1997) for an analysis of the myosin rod]. This may well serve a biological purpose, since a coiled coil the size of myosin might find it difficult to reach its proper register if local coiled-coil interactions formed rapidly and randomly along the

1000-residue sequence. Still, it begs the question of how the native coiled coil folds. An intriguing hypothesis was provided by Steinmetz *et al.* (1998) with the *trigger sequence*, a short autonomous helical folding unit favoring oligomerization. They observed that the folding of cortexillin I, whose coiled-coil domain contains 18 heptads, was entirely dependent on the presence of two heptads close to the C-terminal end, which they proposed triggered the folding of the coiled coil. The crystal structure of the cortexillin I coiled coil revealed an abundance of favorable inter- and intrahelical ionic interactions in the trigger region, as well as tight packing interactions in its hydrophobic core (Burkhard *et al.*, 1998). The authors identified potential trigger sequences in several other coiled coils, based on sequence similarity to the cortexillin trigger, and successfully tested the putative trigger they identified in GCN4 (Kammerer *et al.*, 1998). Since then, trigger sequences were identified in other proteins, such as the macrophage scavenger receptor (Frank *et al.*, 2000), intermediate filament proteins (Wu *et al.*, 2000), hantavirus nucleocapsid protein (Alfadhli *et al.*, 2002), and tropomyosin (Araya *et al.*, 2002). However, the trigger sequences in the various proteins show considerable diversity; it is therefore unlikely that a consensus sequence for trigger sites exists. Rather, as shown by Lee *et al.* (2001) for a cortexillin/GCN4 hybrid lacking the trigger sites, folding could be obtained by a combination of stabilizing mutations that improved helicity, electrostatic interactions, and hydrophobic packing, but did not bring the final sequence closer to the trigger consensus. Thus, trigger sequences may correspond to short regions in coiled coils, which can serve as nucleation sites for folding due to their unusual α-helical stability and high number of interactions stabilizing the oligomeric form. They will only adhere to a consensus as far as the principles guiding coiled-coil stability can be reduced to a consensus sequence.

The main properties that provide stability to a coiled coil are helical propensity, hydrophobicity of the core, tightness of the core packing, shielding of the core from solvent, and favorable polar and ionic interactions. Beyond general statements, the detailed contribution of residues to the stability of natural coiled coils is not well understood. Efforts to elucidate the contribution of different residues, such as the host-guest studies of Hodges on the effects of the 20 amino acids at positions *a* and *d* of a model peptide (Tripet *et al.*, 2000; Wagschal *et al.*, 1999), ignore the complementarity of side-chain packing and the importance of sequence context for judging the effects of a substitution. Correspondingly, the results correlate poorly with the residue frequencies observed in natural coiled coils. It may be argued, however, based on the experience from engineering stable repetitive proteins (Kohl *et al.*, 2003; Main *et al.*, 2003;

Mosavi *et al.*, 2002), that the consensus sequence embodies the most stable form of that particular fold. In that respect, it is noteworthy that coiled coils in archaea (e.g., prefoldin), which have a greater stability than their homologs in eukaryotes, match consensus matrices of coiled-coil sequences better, even though they did not contribute to them (A. Lupas, unpublished observation).

An effect that clearly enhances the stability of coiled coils is the number of helices in the structure. More helices mean a broader hydrophobic interface, better shielding of the core, and more opportunities for polar and ionic interactions by the involvement of positions *b* and *c*. GCN4-pIL, which forms dimers, is thus less stable than the practically sequence-identical GCN4-pLI, which forms tetramers (Harbury *et al.*, 1993). Similarly, the Fos homodimer is much less stable than the tetramer, obtained by mutation of two core lysine residues to norleucine (Campbell *et al.*, 2002). Two of the most stable coiled coils known to us, tetrabrachion and Sendai virus phosphoprotein, are both tetramers. Denaturation of the tetrabrachion coiled coil requires 70% sulfuric acid, fuming trifluoromethanesulfonic acid, or heating to 130 °C for 30 minutes in 6 M guanidinium hydrochloride (Peters *et al.*, 1995). Sendai virus phosphoprotein withstands extensive deletions, point mutations to proline in core residues, and two-residue insertions that move the hydrophobic seam to the opposite face of the helix (Tarbouriech *et al.*, 2000). In both structures, the interactions between neighboring helices appear to provide such a degree of stability that they can overcompensate for the formation of core cavities. Heretically, both structures contain considerable quantities of water in their ''hydrophobic'' core.

IV. Structural Diversity

A. *Fibers and Zippers*

Historically, coiled coils were identified with long fibrous molecules, from which their structural properties had been determined. Fiber diffraction studies on proteins of the k-m-e-f class were highly successful, initially on dried specimens but later also on native samples (Cohen and Holmes, 1963). However, these proteins turned out to be very difficult to analyze by high-resolution X-ray crystallography for the same reasons that made them so amenable to fiber diffraction—their tendency to aggregate into fibers rather than crystals and the extreme dimensions of their asymmetric units. It took decades to obtain a working structure for tropomyosin [at 15 Å resolution (Phillips *et al.*, 1986); at 9 Å (Whitby *et al.*,

1992); at 7 Å (Whitby and Phillips, 2000); N-terminal fragment at 2 Å (Brown *et al.*, 2001); C-terminal fragment at 2.7 Å (Li *et al.*, 2002)], and a high-resolution structure for the entire molecule still remains to be determined. Arguably, the only k-m-e-f protein with a known high-resolution structure is fibrinogen (Brown *et al.*, 2000; Yang *et al.*, 2000, 2001).

Not surprisingly, the Protein Data Bank primarily offers examples of short coiled coils. The first ones to be determined at high resolution were the three-stranded coiled coil in influenza hemaglutinin (Wilson *et al.*, 1981) and the two-stranded coiled coil in the catabolite gene activator protein CAP (McKay and Steitz, 1981), but the coiled coil in CAP was only recognized a decade later (Nilges and Brunger, 1991). In 1988, when the leucine zipper hypothesis was formulated (Landschulz *et al.*, 1988), the coiled-coil crystal structures in PDB could still be counted on the fingers of one hand. The leucine zipper changed research into coiled coils fundamentally by refocusing the studies from long fibers to a broad range of shorter coiled coils found in globular proteins. The leucine zipper also provided an ideal vehicle for exploring issues in coiled-coil packing, oligomerization, folding, and stability, which were discussed in the previous section and, in greater detail, in Chapter 4 by Woolfson. Originally, the leucine zipper was proposed to be an antiparallel arrangement of two helices, which dimerized by the ridges-into-grooves meshing of four leucines in a heptad spacing (Landschulz *et al.*, 1988), but it was soon recognized as a new short form of the coiled coil, in which the stability was derived from the optimized packing of the leucines in position *d* (O'Shea *et al.*, 1991). Since then, the number of coiled-coil structures in PDB has increased dramatically. The two main groups recognizable today are oligomeric structures, which conform to the original ideas about coiled coils, and single-chain antiparallel helical bundles, which were only gradually recognized to form a valid subgroup of the coiled coil proteins (Cohen and Parry, 1986, 1990) (Fig. 1). The leucine zipper can now be seen to belong to a much larger group of short, dimeric coiled coils, parallel and antiparallel, which serve to oligomerize and position many of the most common DNA-binding domains (Fig. 6c), including the basic region, helix turn helix (HtH) domains (such as in CAP), and zinc fingers.

B. Tubes, Sheets, Spirals, Funnels, and Rings

Coiled-coil helices usually have one stripe of residues engaged in knobs-into-holes interactions. There are, however, instances where a helix may engage in such interactions along two stripes (Cohen and Parry, 1990; Walshaw and Woolfson, 2003; Walshaw *et al.*, 2001). For example, helices of four-stranded coiled coils make their primary knobs-into-holes contacts

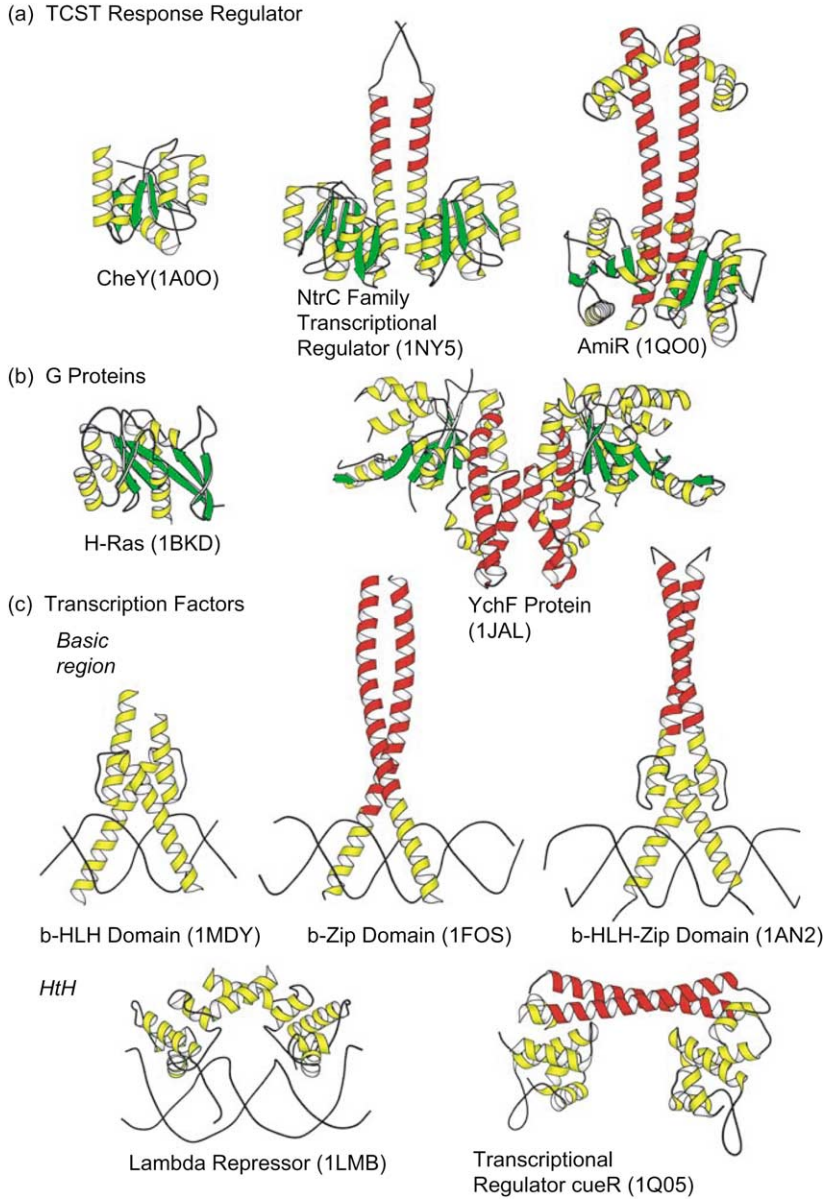


FIG. 6. Coiled coils that ''grow'' out of other folds. (a) Response regulators of two-component signal transduction (TCST) systems dimerizing through the extension of helix $\alpha 5$. (b) Dimerization interface in a GTP-binding protein by extension of helix $\alpha 4$.

with their neighboring helices, using positions *a* and *e* on one side and *d* and *g* on the other. The diagonally opposite helix is furthest away and makes the least number of contacts, frequently to the point where the center of the coiled coil contains cavities. These are sometimes filled by water or other molecules, giving the coiled coil the nature of a tube (Figs. 5 and 7). The tube-shaped character of such coiled coils is naturally enhanced as the number of helices grows, and with it the diameter of the tube. Figure 7 shows examples of tube structures with 4, 5, 6, and 12 helices, the last forming an opening with a diameter of about 20 Å.

As the position of the two stripes diverges on the helix surface towards opposite sides, the association takes the form of a sheet. The structure database offers multiple examples of such sheets, usually containing three or four helices (Fig. 8). As can be seen from their cross sections, forming knobs-into-holes interactions on opposite sides of one helix leads to ''supercoil schizophrenia,'' since the interactions require supercoiling in opposite directions. This results in a straight conformation of the central helices and an increased distortion of the flanking helices. In turn, this means that the helices gradually move out of register, so that a regular packing can only be achieved approximately and over a limited distance.

Sheets meet tubes in the formation of spirals. For example, the major coat protein subunits of filamentous bacteriophage assemble into curved sheets that further assemble into a multistranded spiral (Fig. 9). Coat proteins from different bacteriophages differ in the extent to which they engage in knobs-into-holes interactions and in the degree of curvature of the sheets. Correspondingly different assemblies use different numbers of sheets to form the spiral (five in phage Ike and six in phage Pf1). Similar principles guide the assembly of coiled-coil spirals in two types of bacterial surface structures, flagella and pili. For the flagella of *Salmonella typhimurium*, Namba and colleagues determined the crystal structure of the flagellin subunit (Samatey *et al.*, 2001) and then used electron cryomicroscopy to place the subunits into the R form of the filament (Yonekura *et al.*, 2003); Tainer and colleagues used a similar approach to image the type IV pili of *Vibrio cholerae* and *Pseudomonas aeruginosa* (Craig *et al.*, 2003). A special case of a spiral is given by the multidrug resistance protein MexA, whose crystal structure shows an unusual, hourglass assembly of two spiral arcs packed end to end (Fig. 9). MexA belongs to a class of periplasmic

(c) Coiled-coil mediated dimerization in transcription factors: the basic region-helix loop helix (bHLH) DNA-binding domain (from MyoD) and coiled coils obtained by extending either the basic region helices (from Fos) or the helices of the HLH motif (from Max). The helix-turn-helix DNA-binding domain (from lambda repressor) and its elaboration with a coiled coil (in cueR).

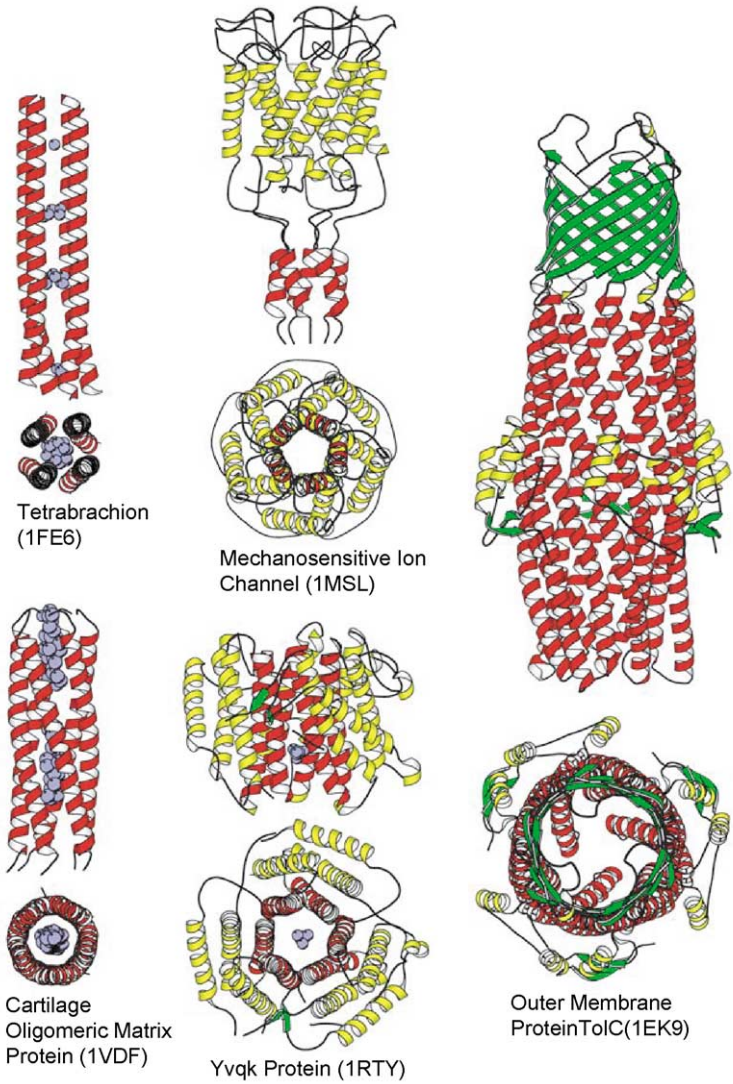


FIG. 7. Coiled-coil tubes. The structures are shown in side view and cross section, and molecules found in the channels are shown in space-filling representation.

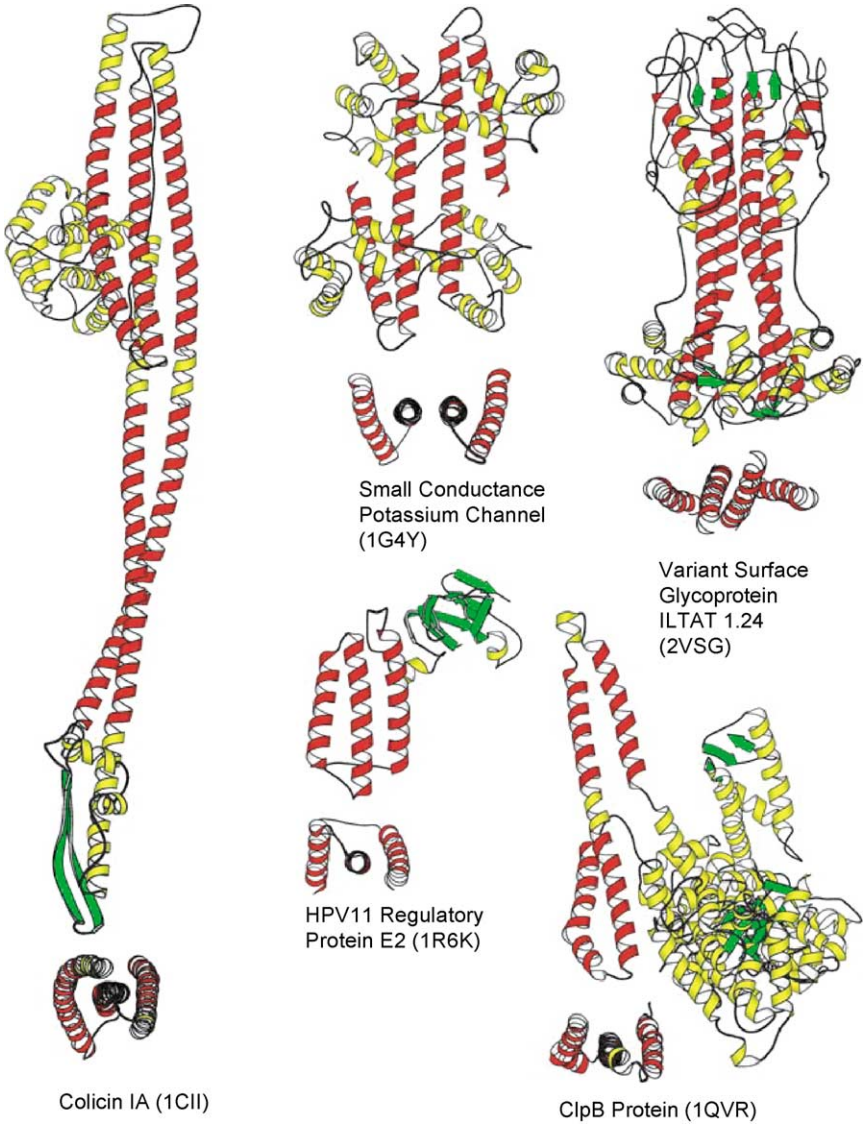


FIG. 8. Coiled-coil sheets. The cross-section of each sheet is shown under the structure.

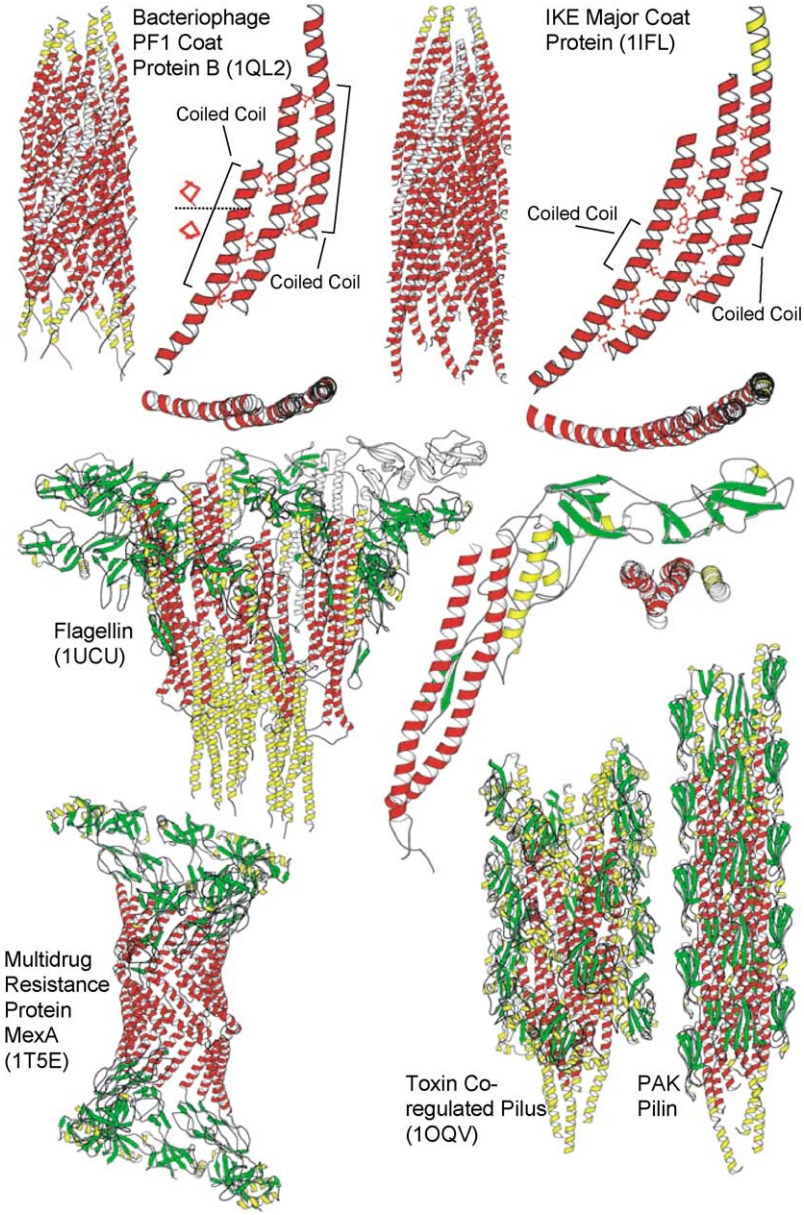


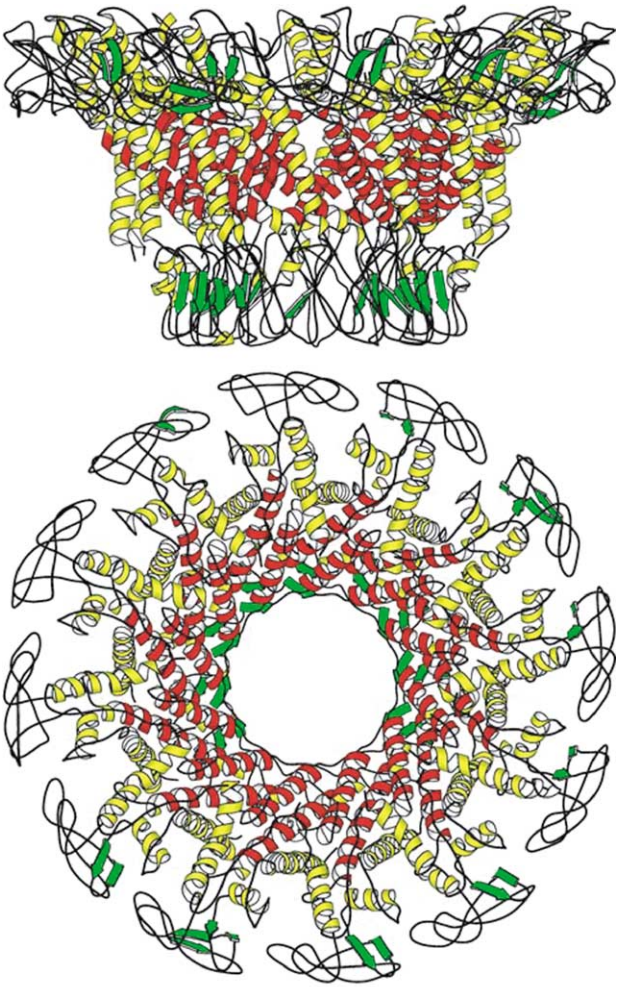
FIG. 9. Coiled-coil spirals. For the phage coat proteins and flagellin, subunits are shown enlarged next to the structures, as well as the cross sections of the coiled-coil sheets they form. The positions of the subunits in the structures are indicated in white. The core packing layers are also shown for the phage coat proteins in order to illustrate the use of knobs-into-holes and ridges-into-grooves layers.

adaptor molecules that connect inner-membrane transporters to outer membrane efflux pores, such as TolC (Fig. 7), yielding an efflux pump complex that spans the entire envelope of Gram-negative bacteria. For biological reasons, this structure is thought to be an artifact of crystallization, and the *in vivo* structure is thought to resemble a funnel (Higgins *et al.*, 2004). Another funnel-shaped assembly based on coiled-coil interactions is formed by the upper collar protein of *Bacillus* bacteriophage Φ29, which is part of the motor that translocates the phage DNA into the precursor capsid (Fig. 10).

A protein that closes the circle back to fibers, from which the coiled-coil structure took its start, is apolipoprotein A-II (Fig.11), an astonishing complex formed by subunits of 77 residues with an underlying 22-residue periodicity. Four subunits assemble into a highly irregular, parallel bundle, which shows knobs-into-holes interactions only locally and only in pairwise helical contacts (but has the supercoil angle expected for a periodicity of 3.67). The helices have multiple symmetry breaks, including a most pronounced one in a central area that unfolds to different degrees in the ''inner'' and ''outer'' subunits, thus bending the bundle. The bent bundles assemble head-to-tail into a spiral-shaped fiber. Despite a homologous origin and a common underlying 22/6 periodicity (Boguski *et al.*, 1985), apolipoproteins show an astounding structural diversity: Apolipoprotein A-I is built of 251-residue subunits that also assemble into long four-stranded helical bundles with limited knobs-into-holes interactions and a right-handed supercoil twist, albeit in an antiparallel orientation and in a staggered way, such that no two N-termini are in the same register. This arrangement results in the formation of a single twisted ring (Fig. 11). In contrast, apolipoprotein E N-terminal domain (191 residues) and the insect apolipophorin III (180 residues) form monomeric, single-chain helical bundles with a left-handed supercoil twist and fairly regular knobs-into-holes interactions (Fig. 11).

C. What Exactly is a Coiled Coil?

The diversity of coiled coils presented in this Chapter raises for us the question: can one hope to distinguish unambiguously every such structure from all others? *A priori*, the peculiar and highly regular structure implied by the standard model would suggest that the answer should be yes; knobs-into-holes packing, supercoiling, symmetric arrangement of the helices, low crossing angle, length of the structure, and distinctness from other domains should provide multiple criteria by which to judge whether a bundle of helices forms a coiled coil. In practice, there is a continuum of structures that stretch along an imaginary axis of ''coiled-coiledness'' and



Bacteriophage ϕ Upper
Collar Protein (1FOU)

FIG. 10. A coiled-coil funnel.

for any criterion, protein domains can be found that violate it and yet would be accepted in general as coiled coils. Regarding distinctness from other domains, we have compiled examples of protein families in Fig. 6, in which coiled-coil domains ''grow'' through the extension of helices that are integral components of other folds, such as helix $\alpha 5$ of the

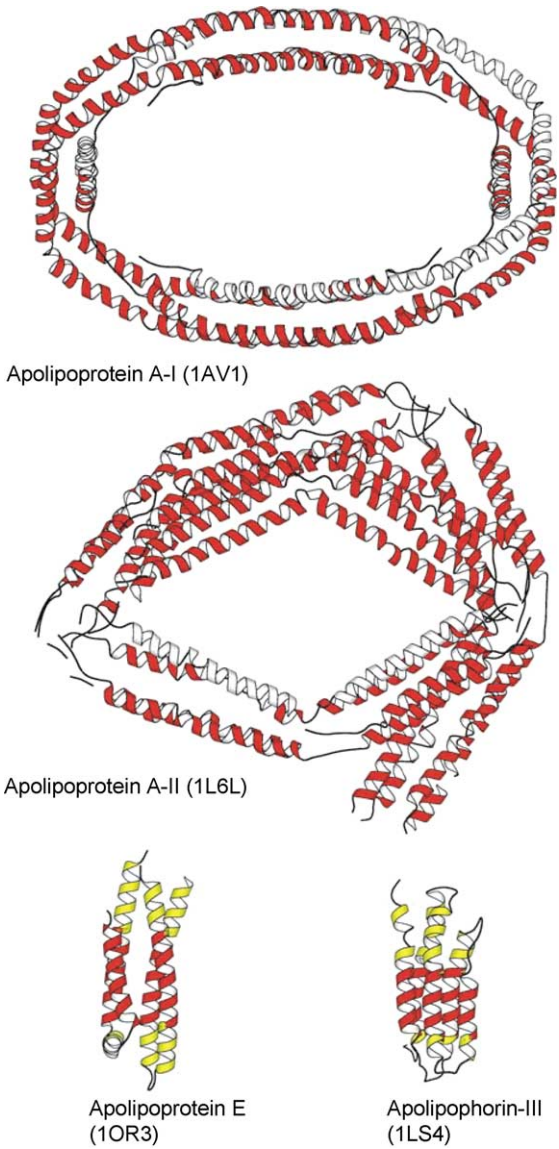


FIG. 11. Structure diversity in apolipoproteins. Despite being of homologous origin and sharing an underlying 22-residue sequence periodicity, different apolipoproteins show fundamental differences in oligomerization state, subunit orientation, and supercoil angle.

flavodoxin fold in bacterial response regulators (Fig. 6a) or helix $\alpha 4$ of the P-loop NTPase fold in YchF (Fig. 6b). In Fig. 12, coiled coils arise within the folds themselves, for example in the flavodoxin-like N-terminal domain of formyl-CoA transferase (1P5H), which dimerizes via coiled-coil interactions in helix $\alpha 5$.

Regarding length of the structure, it is certainly the case that many classical coiled coils are very long and that even the leucine zipper, once considered to mark the short end of the coiled-coil distribution, is four heptads long. There are however several rather clear coiled coils, which are only ten or so residues long, such as in Gal4 (1D66; 14 residues) or the previously discussed homing endonucleases (e.g., 1G9Z; 12 residues). Regarding the crossing angle, many ''square'' helical bundles (following the nomenclature of Harris *et al.*, 1994) have the same crossing angle as canonical coiled coils (around 20 degrees) but show primarily ridges-into-grooves interactions, as do most helical bundles found in the membrane. Conversely, nonheptad coiled coils show a diversity of crossing angles, from about 25 degrees to −25 degrees. Regarding a symmetric arrangement of helices, we have already encountered in this Chapter many examples of symmetry breaks, and the same may be said of supercoiling and the pitch diversity of coiled coils.

This brings us to the one property that is usually regarded as the hallmark of coiled-coil structure: knobs-into-holes packing. As we have shown earlier, knobs-to-knobs interactions are an integral part of nonheptad coiled coils. One may however still judge that coiled coils differ from other structures by the fact that their core layers move in register, rather than showing the staggered arrangements of ridges-into-grooves. Here again, though, there are many exceptions. In spectrin (Fig. 2b), two helices form regular knobs-into-holes interactions, while the third is out of register. In GrpE (Fig. 4), a parallel, homodimeric coiled coil moves out of register after a stutter to show local ridges-into-grooves interactions, then returns to knobs-into-holes packing. In apolipoproteins A-I and A-II, ridges-into-grooves coexist with knobs-into-holes to produce some highly unusual tetrameric coiled coils. Or *are* these still coiled coils? We find that there is no criterion or set of criteria that would allow us to resolve that question.

V. Function Follows Structure

More clearly than in many other protein folds, the function of coiled coils follows from their main structural properties. Coiled coils are usually long, rigid oligomers of helices with regular packing interactions and extended exposed surfaces. This enables them to assemble into large,

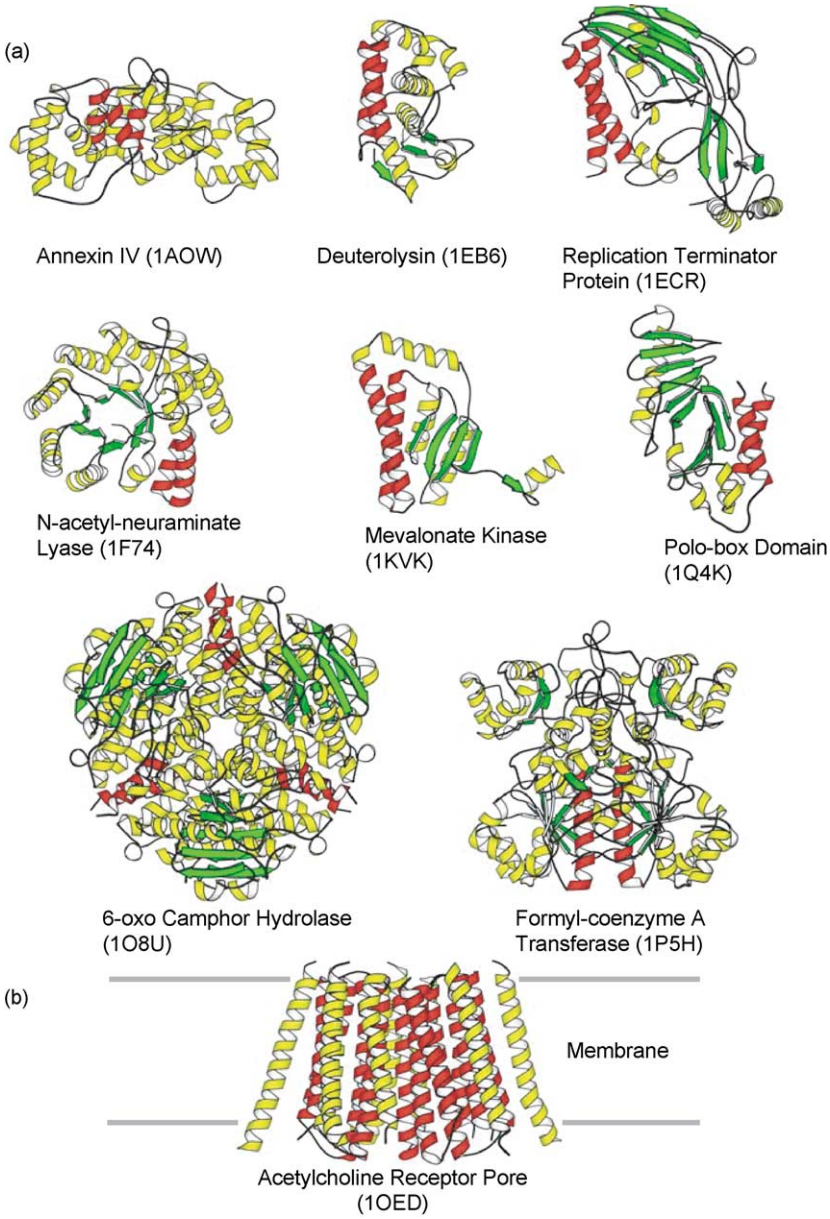


FIG. 12. Coiled coils arising between helices that are part of different folds. (a) Soluble proteins. (b) Example of a membrane protein. Inner membrane proteins are α-helical proteins with an up-and-down topology; their helices therefore favor low crossing angles and frequently show mixtures of knobs-into-holes and ridges-into-grooves interactions.

mechanically rigid structures such as hair, horn, feathers (keratin), and blood clots (fibrinogen); extracellular matrices (laminin) and cytoskeletal networks (intermediate filaments); and a broad array of filaments (flagellins, pilins, phage coat proteins). They are an efficient way to project across large distances and are therefore often found in molecular spacers (murein lipoprotein, Omp α) and the stalks of surface proteins (nonfimbrial adhesins, viral fusion proteins). Their rigidity and large exposed surfaces allow them to function as the arms of proteins involved in the manipulation of polynucleotides (Seryl-tRNA synthetase, DNA polymerase I) and of unfolded polypeptides (ClpB, prefoldin). These arms are single-chain, antiparallel hairpins of helices that, by lacking sequence symmetry, can adapt more precisely to the interaction with asymmetric ligands and can move more readily around a hinge (in a homooligomer, any such motion would involve a symmetry break) (Martin *et al.*, 2004). Above all, coiled coils are ideal mediators of oligomerization, assuming this role in a multitude of proteins such as transcription factors (leucine zippers), molecular motors (myosin, kinesin), receptors (macrophage scavenger receptor, chemotaxis receptors), and signaling molecules (G protein $\beta\gamma$). Their regular packing interactions make these interchangeable to a much larger extent than in most proteins. Correspondingly, they may change their oligomer state and conformation in response to changes in their environment, an ability that is exploited most impressively in membrane fusion proteins, as illustrated by the structures of influenza hemaglutinin at neutral (2HMG) and acidic (1HTM) pH.

In summary, coiled coils are versatile structural elements that are widely used in many protein families. The wealth of forms discovered by high-resolution structure studies in recent years, some of which are truly astonishing even to a seasoned coiled-coil watcher, raise the hope of many more that still remain to be discovered.

References

Alfadhli, A., Steel, E., Finlay, L., Bächinger, H. P., and Barklis, E. (2002). Hantavirus nucleocapsid protein coiled-coil domains. *J. Biol. Chem.* **277,** 27103–27108.

Araya, E., Berthier, C., Kim, E., Yeung, T., Wang, X., and Helfman, D. M. (2002). Regulation of coiled-coil assembly in tropomyosins. *J. Struct. Biol.* **137,** 176–183.

Arndt, K. M., Pelletier, J. N., Müller, K. M., Alber, T., Michnick, S. W., and Plückthun, A. (2000). A heterodimeric coiled-coil peptide pair selected *in vivo* from a designed library-versus-library ensemble. *J. Mol. Biol.* **295,** 627–639.

Arndt, K. M., Pelletier, J. N., Müller, K. M., Plückthun, A., and Alber, T. (2002). Comparison of *in vivo* selection and rational design of heterodimeric coiled coils. *Structure* **10,** 1235–1248.

Arnott, S., and Dover, S. D. (1967). Refinement of bond angles of an α-helix. *J. Mol. Biol.* **30,** 209–212.

Beck, K., Dixon, T. W., Engel, J., and Parry, D. A. D. (1993). Ionic interactions in the coiled-coil domain of laminin determine the specificity of chain assembly. *J. Mol. Biol.* **231,** 311–323.

Beck, K., Gambee, J. E., Kamawal, A., and Bächinger, H. P. (1997). A single amino acid can switch the oligomerization state of the α-helical coiled-coil domain of cartilage matrix protein. *EMBO J.* **16,** 3767–3777.

Berger, B., and Singh, M. (1997). An iterative method for improved protein structural motif recognition. *J. Comput. Biol.* **4,** 261–273.

Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl. Acad. Sci.* **92,** 8259–8263.

Boguski, M. S., Elshourbagy, N., Taylor, J. M., and Gordon, J. I. (1985). Comparative analysis of repeated sequences in rat apolipoproteins A-I, A-IV, and E. *Proc. Natl. Acad. Sci.* **82,** 992–996.

Brown, J. H., Cohen, C., and Parry, D. A. D. (1996). Heptad breaks in α-helical coiled coils: Stutters and stammers. *Proteins* **26,** 134–145.

Brown, J. H., Volkmann, N., Jun, G., Henschen-Edman, A. H., and Cohen, C. (2000). The crystal structure of modified bovine fibrinogen. *Proc. Natl. Acad. Sci.* **97,** 85–90.

Brown, J. H., Kim, K. H., Jun, G., Greenfield, N. J., Dominguez, R., Volkmann, N., Hitchcock-DeGregori, S. E., and Cohen, C. (2001). Deciphering the design of the tropomyosin molecule. *Proc. Natl. Acad. Sci.* **98,** 8496–8501.

Burkhard, P., Steinmetz, M. O., Schulthess, T., Landwehr, R., Aebi, U., and Kammerer, R. A. (1998). Crystallization and preliminary X-ray diffraction analysis of the 190-A-long coiled-coil dimerization domain of the actin-bundling protein cortexillin I from *Dictyostelium discoideum. J. Struct. Biol.* **122,** 293–296.

Campbell, K. M., Sholders, A. J., and Lumb, K. J. (2002). Contribution of buried lysine residues to the oligomerization specificity and stability of the Fos coiled coil. *Biochemistry* **41,** 4866–4871.

Chothia, C., Levitt, M., and Richardson, D. (1977). Structure of proteins: Packing of α-helices and pleated sheets. *Proc. Natl. Acad. Sci.* **74,** 4130–4134.

Chothia, C., Levitt, M., and Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol.* **145,** 215–250.

Cohen, C., and Holmes, K. C. (1963). X-ray diffraction evidence for α-helical coiled-coils in native muscle. *J. Mol. Biol.* **6,** 423–432.

Cohen, C., and Parry, D. A. D. (1986). α-Helical coiled coils—A widespread motif in proteins. *Trends Biochem. Sci.* **11,** 245–248.

Cohen, C., and Parry, D. A. D. (1990). α-Helical coiled coils and bundles: How to design an α-helical protein. *Proteins* **7,** 1–15.

Conway, J. F., and Parry, D. A. D. (1990). Structural features in the heptad substructure and longer range repeats of two-stranded α-fibrous proteins. *Int. J. Biol. Macromol.* **12,** 328–334.

Conway, J. F., and Parry, D. A. D. (1991). Three-stranded α-fibrous proteins: The heptad repeat and its implication for structure. *Int. J. Biol. Macromol.* **13,** 14–16.

Craig, L., Taylor, R. K., Pique, M. E., Adair, B. D., Arvai, A. S., Singh, M., Lloyd, S. J., Shin, D. S., Getzoff, E. D., Yeager, M., Forest, K. T., and Tainer, J. A. (2003). Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol. Cell* **11,** 1139–1150.

Crick, F. H. C. (1952). Is α-keratin a coiled coil? *Nature* **170,** 882–883.

Crick, F. H. C. (1953a). The Fourier transform of a coiled-coil. *Acta Crystallogr.* **6,** 685–689.

Crick, F. H. C. (1953b). The packing of α-helices: Simple coiled-coils. *Acta Crystallogr.* **6,** 689–697.

Delorenzi, M., and Speed, T. (2002). An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics* **18,** 617–625.

Dure, 3rd, L. (1993). A repeating 11-mer amino acid motif and plant desiccation. *Plant J.* **3,** 363–369.

Fong, J. H., Keating, A. E., and Singh, M. (2004). Predicting specificity in bZIP coiled-coil protein interactions. *Genome Biol.* **5,** R11.

Frank, S., Lustig, A., Schulthess, T., Engel, J., and Kammerer, R. A. (2000). A distinct seven-residue trigger sequence is indispensable for proper coiled-coil formation of the human macrophage scavenger receptor oligomerization domain. *J. Biol. Chem.* **275,** 11672–11677.

Fraser, R. D. B., and MacRae, T. P. (1973). *In* ''Conformation in Fibrous Proteins and Related Synthetic Polypeptides,'' pp. 456–465. Academic Press, London.

Glover, J. N., and Harrison, S. C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373,** 257–261.

Gonzales, L., Jr., Brown, R. A., Richardson, D., and Alber, T. (1996a). Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. *Nat. Struct. Biol.* **3,** 1002–1010.

Gonzales, L., Jr., Woolfson, D. N., and Alber, T. (1996b). Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* **3,** 1011–1018.

Gruber, M., and Lupas, A. N. (2003). Historical review: Another 50th anniversary—New periodicities in coiled coils. *Trends Biochem. Sci.* **28,** 679–685.

Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262,** 1401–1407.

Harbury, P. B., Kim, P. S., and Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature* **371,** 80–83.

Harbury, P. B., Plecs, J. J., Tidor, B., Alber, T., and Kim, P. S. (1998). High-resolution protein design with backbone freedom. *Science* **282,** 1462–1467.

Harris, N. L., Presnell, S. R., and Cohen, F. E. (1994). Four helix bundle diversity in globular proteins. *J. Mol. Biol.* **236,** 1356–1368.

Hicks, M. R., Holberton, D. V., Kowalczyk, C., and Woolfson, D. N. (1997). Coiled-coil assembly by peptides with non-heptad sequence motifs. *Fold. Des.* **2,** 149–158.

Higgins, M. K., Bokma, E., Koronakis, E., Hughes, C., and Koronakis, V. (2004). Structure of the periplasmic component of a bacterial drug efflux pump. *Proc. Natl. Acad. Sci.* **101,** 9994–9999.

Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A., and Heesemann, J. (2000). Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J.* **19,** 5989–5999.

Holberton, D., Baker, D. A., and Marshall, J. (1988). Segmented α-helical coiled-coil structure of the protein giardin from the *Giardia* cytoskeleton. *J. Mol. Biol.* **204,** 789–795.

Holton, J., and Alber, T. (2004). Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci.* **101,** 1537–1542.

Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996). High resolution NMR solution structure of the leucine zipper domain of the c-Jun homodimer. *J. Biol. Chem.* **271,** 13663–13667.

Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aebi, U., and Steinmetz, M. O. (1998). An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl. Acad. Sci.* **95,** 13419–13424.

Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Plückthun, A., and Grutter, M. G. (2003). Designed to be stable: Crystal structure of a consensus ankyrin repeat protein. *Proc. Natl. Acad. Sci.* **100,** 1700–1705.

Krylov, D., Mikhailenko, I., and Vinson, C. (1994). A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J.* **13,** 2849–2861.

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240,** 1759–1764.

Lee, D. L., Lavigne, P., and Hodges, R. S. (2001). Are trigger sequences essential in the folding of two-stranded α-helical coiled-coils? *J. Mol. Biol.* **306,** 539–553.

Li, Y., Mui, S., Brown, J. H., Strand, J., Reshetnikova, L., Tobacman, L. S., and Cohen, C. (2002). The crystal structure of the C-terminal fragment of striated-muscle α-tropomyosin reveals a key troponin T recognition site. *Proc. Natl. Acad. Sci.* **99,** 7378–7383.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., and Eisenberg, D. (1993). Crystal structure of a synthetic triple-stranded α-helical bundle. *Science* **259,** 1288–1293.

Lupas, A. (1996a). Prediction and analysis of coiled-coil structures. *Meth. Enzym.* **266,** 513–525.

Lupas, A. (1996b). Coiled coils: New structures and new functions. *Trends Biochem. Sci.* **21,** 375–382.

Lupas, A., Van Dyke, S., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252,** 1162–1164.

Lupas, A., Müller, S., Goldie, K., Engel, A. M., Engel, A., and Baumeister, W. (1995). Model structure of the Ompα rod, a parallel four-stranded coiled coil from the hyperthermophilic eubacterium *Thermotoga maritima*. *J. Mol. Biol.* **248,** 180–189.

Main, E. R., Xiong, Y., Cocco, M. J., D'Andrea, L., and Regan, L. (2003). Design of stable α-helical arrays from an idealized TPR motif. *Structure* **11,** 497–508.

Marshall, J., and Holberton, D. V. (1993). Sequence and structure of a new coiled coil protein from a microtubule bundle in *Giardia*. *J. Mol. Biol.* **231,** 521–530.

Marshall, J., and Holberton, D. V. (1995). *Giardia* gene predicts a 183 kDa nucleotide-binding head-stalk protein. *J. Cell Sci.* **108,** 2683–2692.

Martin, J., Gruber, M., and Lupas, A. N. (2004). Coiled coils meet the chaperone world. *Trends Biochem. Sci.* **29,** 455–458.

McKay, D. B., and Steitz, T. A. (1981). Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290,** 744–749.

McLachlan, A. D., and Karn, J. (1983). Periodic features in the amino acid sequence of nematode myosin rod. *J. Mol. Biol.* **164,** 605–626.

McLachlan, A. D., and Stewart, M. (1975). Tropomyosin coiled-coil interactions: Evidence for an unstaggered structure. *J. Mol. Biol.* **98,** 293–304.

McLachlan, A. D., and Stewart, M. (1976). The 14-fold periodicity in α-tropomyosin and the interaction with actin. *J. Mol. Biol.* **103,** 271–298.

Mittl, P. R. E., Deillon, C., Sargent, D., Liu, N., Klauser, S., Thomas, R. M., Gutte, B., and Grütter, M. G. (2000). The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. *Proc. Natl. Acad. Sci.* **97,** 2562–2566.

Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M., and Hodges, R.S. (1996). Formation of parallel and antiparallel coiled-coils controlled by the relative positions of alanine residues in the hydrophobic core. *J. Biol. Chem.* **271,** 3995–4001.

Mosavi, L. K., Minor, D. L., Jr., and Peng, Z. Y. (2002). Consensus-derived structural determinants of the ankyrin repeat motif. *Proc. Natl. Acad. Sci.* **99,** 16029–16034.

Nilges, M., and Brunger, A. T. (1991). Automated modeling of coiled coils: Application to the GCN4 dimerization region. *Protein Eng.* **4,** 649–659.

Nooren, I. M. A., Kaptein, R., Sauer, R. T., and Boelens, R. (1999). The tetramerization domain of the Mnt repressor consists of two right-handed coiled coils. *Nat. Struct. Biol.* **6,** 755–759.

O'Donoghue, S. I., and Nilges, M. (1997). Tertiary structure prediction using mean-force potentials and internal energy functions: Successful prediction for coiled-coil geometries. *Fold. Des.* **2,** 47–52.

Offer, G., and Sessions, R. (1995). Computer modelling of the α-helical coiled coil: Packing of side-chains in the inner core. *J. Mol. Biol.* **249,** 967–987.

Offer, G., Hicks, M. R., and Woolfson, D.N. (2002). Generalized Crick equations for modeling noncanonical coiled coils. *J. Struct. Biol.* **137,** 41–53.

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254,** 539–544.

Parry, D. A. D. (1975). Analysis of the primary sequence of α-tropomyosin from rabbit skeletal muscle. *J. Mol. Biol.* **98,** 519–535.

Parry, D. A. D. (1982). Coiled-coils in α-helix-containing proteins: Analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Biosci. Rep.* **2,** 1017–1024.

Parry, D. A. D., Crewther, W. G., Fraser, R. D., and MacRae, T. P. (1977). Structure of α-keratin: Structural implication of the amino acid sequences of the type I and type II chain segments. *J. Mol. Biol.* **113,** 449–454.

Pauling, L., and Corey, R. B. (1950). Two hydrogen-bonded spiral configurations of the polypeptide chain. *J. Am. Chem. Soc.* **72,** 534.

Pauling, L., and Corey, R. B. (1953). Compound helical configurations of polypeptide chains: Structure of proteins of the α-keratin type. *Nature* **171,** 59–61.

Pauling, L., Corey, R. B., and Branson, H. R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Chemistry* **37,** 205–211.

Perutz, M. F. (1951). The 1-5-A reflexion from proteins and polypeptides. *Nature* **168,** 653–654.

Peters, J., Nitsch, M., Kuhlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., Engelhardt, H., Pfander, J. P., Müller, S., Goldie, K., Engel, A., Stetter, K.-O., and Baumeister, W. (1995). Tetrabrachion: A filamentous archaebacterial surface protein assembly of unusual structure and extreme stability. *J. Mol. Biol.* **245,** 385–401.

Peters, J., Baumeister, W., and Lupas, A. (1996). Hyperthermostable surface layer protein tetrabrachion from the archaebacterium *Staphylothermus marinus*: Evidence for the presence of a right-handed coiled coil derived from the primary structure. *J. Mol. Biol.* **257,** 1031–1041.

Phillips, G. N., Jr. (1992). What is the pitch of the α-helical coiled coil? *Proteins* **14,** 425–429. Erratum in *Proteins* **17**, 220.

Phillips, G. N., Jr., Fillers, J. P., and Cohen, C. (1986). Tropomyosin crystal structure and muscle regulation. *J. Mol. Biol.* **192,** 111–131.

Samatey, F. A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M., and Namba, K. (2001). Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature* **410,** 331–337.

Seo, J., and Cohen, C. (1993). Pitch diversity in α-helical coiled coils. *Proteins* **15,** 223–234. Erratum in *Proteins* **17**, 219.

Singh, M., Berger, B., Kim, P. S., Berger, J. M., and Cochran, A. G. (1998). Computational learning reveals coiled-coil like motifs in histidine kinase linker domains. *Proc. Natl. Acad. Sci.* **95,** 2738–2743.

Singh, M., Berger, B., and Kim, P. S. (1999). LearnCoil-VMF: Computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins. *J. Mol. Biol.* **290,** 1031–1041.

Steinmetz, M. O., Stock, A., Schulthess, T., Landwehr, R., Lustig, A., Faix, J., Gerisch, G., Aebi, U., and Kammerer, R. A. (1998). A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *EMBO J.* **17,** 1883–1891.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R. A. (2000). Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–776.

Strelkov, S. V., and Burkhard, P. (2002). Analysis of α-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J. Struct. Biol.* **137,** 54–64.

Tarbouriech, N., Curran, J., Ruigrok, R. W. H., and Burmeister, W. P. (2000). Tetrameric coiled coil domain of Sendai virus phosphoprotein. *Nat. Struct. Biol.* **7,** 777–781.

Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000). Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d." *J. Mol. Biol.* **300,** 377–402.

Trybus, K. M., Freyzon, Y., Faust, L. Z., and Sweeney, H. L. (1997). Spare the rod, spoil the regulation: Necessity for a myosin rod. *Proc. Natl. Acad. Sci.* **94,** 48–52.

Vinson, C. R., Hai, T., and Boyd, S. M. (1993). Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: Prediction and rational design. *Genes Dev.* **7,** 1047–1058.

Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999). The role of position *a* in determining the stability and oligomerization state of α-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci.* **8,** 2312–2329.

Walshaw, J., and Woolfson, D. N. (2001). SOCKET: A program for identifying and analyzing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307,** 1427–1450.

Walshaw, J., and Woolfson, D. N. (2003). Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J. Struct. Biol.* **144,** 349–361.

Walshaw, J., Shipway, J. M., and Woolfson, D. N. (2001). Guidelines for the assembly of novel coiled-coil structures: α-Sheets and α-cylinders. *Biochem. Soc. Symp.* **68,** 111–123.

Whitby, F. G., and Phillips, G. N., Jr. (2000). Crystal structure of tropomyosin at 7 Å resolution. *Proteins* **38,** 49–59.

Whitby, F. G., Kent, H., Stewart, F., Stewart, M., Xie, X., Hatch, V., Cohen, C., and Phillips, G. N., Jr. (1992). Structure of tropomyosin at 9 Å resolution. *J. Mol. Biol.* **227,** 441–452.

Wilson, I. A., Shekhel, J. J., and Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289,** 366–373.

Wolf, E., Kim, P. S., and Berger, B. (1997). MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Science* **6,** 1179–1189.

Woolfson, D. N., and Alber, T. (1995). Predicting oligomerization states of coiled coils. *Protein Sci.* **4,** 1596–1607.

Wu, K. C., Bryan, J. T., Morasso, M. I., Jang, S.-I., Lee, J.-H., Yang, J.-M., Marekov, L. N., Parry, D. A. D., and Steinert, P. M. (2000). Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. *Mol. Biol. Cell* **11,** 3539–3558.

Yang, Z., Mochalkin, I., Veerapandian, L., Riley, M., and Doolittle, R.F. (2000). Crystal structure of native chicken fibrinogen at 5.5 Å resolution. *Proc. Natl. Acad. Sci.* **97,** 3907–3912.

Yang, Z., Kollman, J. M., Pandi, L., and Doolittle, R. F. (2001). Crystal structure of native chicken fibrinogen at 2.7 Å resolution. *Biochemistry* **40,** 12515–12523.

Yonekura, K., Maki-Yonekura, S., and Namba, K. (2003). Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424,** 643–650.

Zhang, L., and Hermans, J. (1993). Calculation of the pitch of the α-helical coiled coil: An addendum. *Proteins* **17,** 217–218.

THE DESIGN OF COILED-COIL STRUCTURES AND ASSEMBLIES

By DEREK N. WOOLFSON

Department of Biochemistry, School of Life Sciences, University of Sussex, Falmer BN1 9QG, United Kingdom

I. Introduction to Protein Design 80
II. Rules for Coiled-Coil Design: The Basics of Coiled-Coil Sequence and Structure 82
A. Hydrophobic Interactions and Helix Formation 83
B. van der Waals' Forces, Steric Constraints, and Oligomer Specification 86
C. Salt-Bridge Interactions and Partner Selection 90
D. Buried Polar Groups: The Icing on the Cake 91
III. Key Coiled-Coil Designs 92
A. Parallel Structures 92
B. Antiparallel Structures 101
IV. Summary 105
References 106

Abstract

Protein design allows sequence-to-structure relationships in proteins to be examined and, potentially, new protein structures and functions to be made to order. To succeed, however, the protein-design process requires reliable rules that link protein sequence to structure/function. Although our present understanding of coiled-coil folding and assembly is not complete, through numerous bioinformatics and experimental studies there are now sufficient rules to allow confident design attempts of naturally observed and even novel coiled-coil motifs. This review summarizes the current design rules for coiled coils, and describes some of the key successful coiled-coil designs that have been created to date. The designs range from those for relatively straightforward, naturally observed structures—including parallel and antiparallel dimers, trimers and tetramers, all of which have been made as homomers and heteromers—to more exotic structures that expand the repertoire of Nature's coiled-coil structures. Examples in the second bracket include a probe that binds a cancer-associated coiled-coil protein; a tetramer with a right-handed supercoil; sticky-ended coiled coils that self-assemble to form fibers; coiled coils that switch conformational state; a three-component two-stranded coiled coil; and an antiparallel dimer that directs fragment complementation of larger proteins. Some of the more recent examples show an important

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70004-2

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

development in the field; namely, new designs are being created with function as well as structure in mind. This will remain one of the key challenges in coiled-coil design in the next few years. Other challenges that lie ahead include the need to discover more rules for coiled-coil prediction and design, and to implement these in prediction and design algorithms. The considerable success of coiled-coil design so far bodes well for this, however. It is likely that these challenges will be met and surpassed.

I. Introduction to Protein Design

Why bother with protein design? After all, nature has produced a wide variety of beautiful protein structures with fascinating functions. The author's response to this question has three points. First, protein design provides the acid test of our understanding of the informational aspect of the protein-folding problem and, in particular, how protein sequence relates to the three-dimensional structure (and function) of proteins. Second, protein design attempts to capture the salient features of protein structure and function in simpler contexts. In other words, any natural protein sequence contains superimposed information about the protein's folding, structure and stability, and its function(s). Protein design attempts to disentangle and use this information. Third, natural proteins represent only a tiny fraction of the possible protein sequence space available. Whether they also represent a fraction of the possible protein structure space is another question. It is not yet clear whether the natural protein structures observed so far, which suggest a limited number of protein folds (Liu *et al.*, 2004a; Wolf *et al.*, 2000), are in fact significantly greater. Whatever the case, protein design offers possibilities for exploring protein sequences, structures, and functions beyond those examined by nature. In turn, it presents a route to new structures and functions with potentially exploitable applications.

There are a number of approaches to protein engineering and design, which, for the purposes of this review, are defined here. *Protein engineering* in general refers to the process of making one or a relatively small number of mutations in a natural protein framework to examine sequence-to-structure/function relationships in proteins, and/or to improve their structural properties or functions. At the other extreme, *de novo protein design* refers to attempts to construct totally new protein sequences with prescribed structures (and functions) from first principles. By *protein redesign*, this author refers to the process of mutating already designed scaffolds to create new molecules with improved structural properties, stabilities, and functions.

Clearly, in *de novo* design the starting protein sequence is chosen by the designer. In protein engineering and redesign, there are a number of experimental routes to making mutants. These can be made one or a small number at a time, in which case the process is iterative. Alternatively, many mutations can be made simultaneously, either at specific sites in the scaffold (saturation mutagenesis) or randomly throughout the protein (random mutagenesis) to create a library of mutants from which variants with desired properties are selected. These methods are referred to as combinatorial or semirational. Finally, in both the rational and combinatorial approaches, the mutations can be guided by computational methods and searches.

Two important concepts in *de novo* protein design are those of positive and negative design. In positive design, sequence-to-structure rules are used to direct the formation of and stabilize the target structure, whereas negative design refers to the idea of designing against (i.e., destabilizing) alternative and often competing structures (Beasley and Hecht, 1997; Hellinga, 1997; Hill *et al.*, 2000). The application of these principles to coiled-coil structures is particularly important because, at least at first sight, coiled coils share a straightforward repeat sequence and the possibilities for forming the wrong quaternary structure are significant.

So far, all references to designing protein functions have been bracketed. This is because the current state-of-the-art in protein design is at the structural level. In other words, we are better at designing protein structures than we are at introducing function. This is not to say that the redesign and *de novo* design of protein function is totally beyond us. Indeed, protein engineers have been modifying natural protein frameworks and their functions for around two decades.

This review focuses almost exclusively on the rational design and redesign of coiled-coil motifs. Many reviews have been written on protein design in general. Recent papers that, in the view of this author, reflect the current state of the art in globular protein *de novo* design of structure and redesign of function are from Baker (Kuhlman *et al.*, 2003) and Hellinga (Dwyer *et al.*, 2004; Looger *et al.*, 2003), respectively. Discussion of the design and redesign of four-helix bundle proteins is limited. Although some four-helix bundles can also be considered as coiled coils, this is not always the case. Also, with the vagaries of design in the absence of full structure determination, this author considers it best not to group all four-helix bundles together with coiled coils at present. Finally, reviews on the design of four-helix bundles are available (Hill *et al.*, 2000; Woolfson, 2001). Regarding the design of coiled-coil motifs there have been a number of commentaries and review papers on this over the years (Cohen and Parry, 1990, 1994; Kohn and Hodges, 1998; MacPhee and Woolfson,

2004; Micklatcher and Chmielewski, 1999; Oakley and Hollenbeck, 2001; Schneider *et al.*, 1998). Finally, in addition to Chapters 3 and 4 in the volume on coiled-coil sequences and structures, there have been a number of reviews of the coiled-coil field in general (Burkhard *et al.*, 2001; Gruber and Lupas, 2003; Lupas, 1996).

The review is laid out as follows: in Part II, there is a discussion of the structural features of the coiled-coil assemblies and the current state-of-the-art of sequence-to-structure rules pertinent to design; in Part III, there is a detailed description of some of the key coiled-coil designs, including the rules used to create them and their significance in the field; and, finally, there is a discussion of the potential for new coiled-coil designs and where the field might go in the future.

II. Rules for Coiled-Coil Design: The Basics of Coiled-Coil Sequence and Structure

Any design and engineering process requires an understanding of how to assemble the basic building blocks available into the objects being targeted. Though it is not always absolutely necessary, ideally this understanding should be at as fundamental a level as possible. Protein design and assembly are no different. Here, we refer to *rules* that link protein sequence and structure. These rules can be general, such as ''place hydrophobic amino acids alternately three and four residues apart to direct the folding and assembly of amphipathic α-helices,'' or a little more specific, such as ''make every second hydrophobic amino acid Leu to guide the assembly of dimers.'' These are examples of rules of thumb that allow protein designers to build up a protein sequence compatible with a desired target structure. They can be used manually or built into algorithms to be implemented computationally. In many cases, these rules work very well and, if engineering was the only goal of protein design, that might be the end of it. However, it is much more powerful, and ultimately more intellectually satisfying, to understand the fundamental basis of the rules. In terms of protein design, the chemical level of understanding suffices (i.e., how the sequence-to-structure rules can be rationalized in terms of the underlying noncovalent forces). For this reason, the current state-of-the-art of the rules for coiled-coil design is presented below in terms of the noncovalent forces that direct coiled-coil folding and assembly.

For a more thorough description and discussion of coiled-coil structures, the reader is referred to Chapter 3. For the purposes of this review, coiled coils are described as canonical or noncanonical. Canonical coiled coils are those based on tandem heptad sequence repeats that form

right-handed amphipathic α-helices, which then assemble to form helical bundles with left-handed supercoils. In contrast, noncanonical α-helical coiled coils are built from nonheptad-based repeats and, as a consequence, do not necessarily form coiled coils with left-handed or even regular supercoils. For instance, 11-residue repeat sequences give rise to right-handed coiled coils. The majority of this review focuses on designs made using canonical, heptad-based coiled coils as their starting points.

A. *Hydrophobic Interactions and Helix Formation*

As with most, if not all, biological self-assembly processes, a key driving force in coiled-coil folding and assembly is the hydrophobic effect. Put simply, the hydrophobic effect is the phenomenon that hydrocarbon and water tend to phase separate. Thus, when in the context of a biological, aqueous buffer, biological macromolecules fold or self-assemble to minimize the hydrophobic surface area in contact with bulk solvent. In the case of proteins, which can be considered as essentially linear heteropolymers of hydrophobic (*H*: Ala, Phe, Ile, Leu, Met, Val, Trp, and Tyr) and polar (*P*: Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, and Thr) residues, the polypeptide chain folds to bury *H* residues and expose *P* side chains. (In these terms, there is also a third category of amino acid residue, which can be designated as special [*S*: Cys, Gly, and Pro]. These amino acids confer unusual conformational properties on the polypeptide chain. However, for reasons outlined in the main text, these residues occur less frequently in coiled-coil sequences compared with protein sequences as a whole.) Therefore, to a first approximation, the structures arrived at through this reaction to the aqueous medium reflect the pattern of *H* and *P* residues along the primary sequence. For instance, alternating patterns of *H* and *P* residues (*HPHP*...) tend to adopt β-strand conformations and lead to β-sheet-based structures. This is because alternate residues along a β-strand point in opposite directions out from the strand, so an alternating *HP* pattern produces an amphipathic β-strand (i.e., a strand with the *H* residues sequestered on one side and the *P* residues on the other). In turn, two or more such strands come together to form an amphipathic β-sheet, which can pack into a globular structure to bury its hydrophobic face.

As the α-helix has 3.6 residues per turn, hydrophobic side chains spaced at combinations of three and four residues apart are required to make an amphipathic structure. For helices in globular proteins, a variety of combinations of three-plus-four spacings of hydrophobic residues are observed (Chothia *et al.*, 1981). This leads to a range of helix-helix packing angles and arrangements. However, to form more persistent fibrous coiled-coil

structures, more regular *HP* patterns are required. The canonical heptad pattern, *HPPHPPP*, seems to predominate in nature. The full reasons for this are not clear but, as described below, it is likely that it leads to the most efficient way of packing two or more helices in a fibrous helical bundle.

The *HPPHPPP* pattern is usually denoted *abcdefg*, with *a* and *d* assigned to the *H* residues, and can be pictured in terms of a helical wheel (Fig. 1C and D). Tandem heptad repeats along a polypeptide chain give an average separation between *H* residues of 3.5 residues. As this falls short of the 3.6-residues per turn of a regular α-helix, the *a* plus *d* hydrophobic face tracks around the helix with its own helical pitch, which is slighter and in the opposite direction of the backbone α-helix. Thus, when two or more helices pack together to form a coiled-coil oligomer, they do not pack straight as suggested by standard helical wheel diagrams, but wrap around one another in order to maximize contacts between the hydrophobic surfaces (Fig. 1A and D). The sense of this so-called *supercoiling* is the same as the hydrophobic stripe. Thus, in the case of a canonical coiled coil, the helices pack with a left-handed supercoil.

In terms of coiled-coil design several questions emerge from this rudimentary structural analysis. First, what is the minimum number of contiguous heptad repeats required to achieve a stably folded coiled-coil structure? The full answer to this is a little involved because for some natural coiled-coil sequences ''trigger'' sequences are believed to be essential for folding, even for very long sequences (Burkhard *et al.*, 2000a; Kammerer *et al.*, 1998; Lee *et al.*, 2001; Steinmetz *et al.*, 1998; Walshaw and Woolfson, 2001b). However, for design purposes, three or four heptad repeats appear to be sufficient for stable coiled-coil folding (Lumb *et al.*, 1994; Su *et al.*, 1994; Talbot and Hodges, 1982), though, as discussed below, a two-heptad design has been described, albeit with poor oligomer-state specificity (Burkhard *et al.*, 2000b; Meier *et al.*, 2002).

Second, what type of residues are best at the *a* and *d* sites? This is also a difficult question to answer directly because, as will be addressed in more detail over the next few sections, the nature of residues at these sites influences coiled-coil stability, oligomer state, partner selection, and helix-helix orientation (Table I). However, in general terms, natural coiled-coil sequences tend to use the aliphatic hydrophobic residues (Ala, Ile, Leu, Met, and Val) at these positions, rather than the aromatic hydrophobic side chains (Phe, Trp and Tyr) (Parry, 1982; Woolfson and Alber, 1995). The reason for this is probably a combination of bulk and steric constraints presented by the aromatic residues. However, a thorough understanding of the possible exclusion of aromatic side chains from coiled-coil

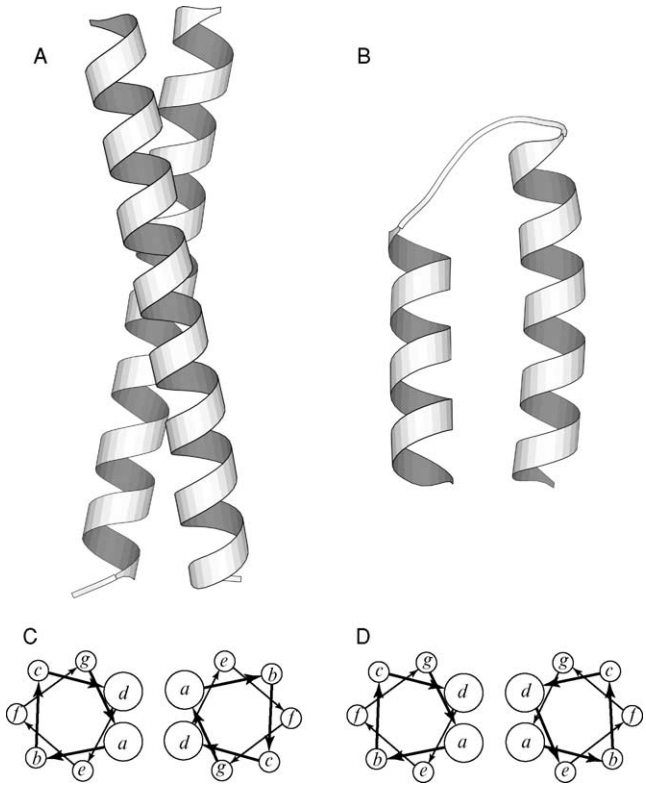


FIG. 1. (A) A dimeric coiled coil [pdb identifier, 2zta; (O'Shea *et al.*, 1991)] (B) An antiparallel, two-stranded coiled coil [pdb identifier, 1sry; (Fujinaga *et al.*, 1993)]. The latter has been cropped to a region of similar length to the leucine-zipper peptide. These figures were generated using Molscript (Kraulis, 1991). (C and D) The heptad-repeat sequence *abcdefg* mapped onto these two types of structure (i.e., parallel and antiparallel two-helix structures, respectively). Reproduced from Pandya *et al.* (2004).

cores awaits further protein-engineering experiments in which multiple aromatic residues are introduced at *a* and *d*. Note added in proof: one such experiment has now been done (see Liu *et al.*, 2004b), and an $a = \mathrm{d} = \mathrm{Trp}$ peptide forms a pentamer.

Third, what residues are best at the remaining *b, c, e, f,* and *g* sites? Again, this will be addressed in more detail below. However, in general, these positions are more permissive than the *a* and *d* sites, though polar and helix-favoring residues (Ala, Glu, Lys, and Gln) tend to be favored both by nature and by protein designers for these positions.

TABLE I
Preferred Amino Acids at the *a* and *d* Sites of Known Coiled-Coil Structures

Structural type	Most preferred amino acids	
	a	*d*
Parallel two-stranded	Asn > Val > Leu > Ile (1.77)	Leu > Met
Parallel three-stranded	Ile > Val > Leu (1.91)	Leu >> Ile
Parallel four-stranded	Met ≈ Ile > Val > Leu > Ala	Met > Ile > Gln
Parallel five-stranded	Leu > Ile	Met > Gln > Thr > Val
Antiparallel two-stranded	Leu ≈ Ile	Leu

These residues occurred twice or more often than expected by chance at the SOCKET-assigned heptad positions of structurally verified coiled-coil motifs in the Protein Data Bank (Walshaw and Woolfson, 2001b). These data are taken from amino acid profiles that were generated by normalizing the raw count data by the observed frequencies of each amino acid in the SWISSPROT database. The bracketed figures after some of the entries give the normalized frequency for that amino acid at that position; in these cases, the frequencies fall below the nominal cutoff of 2 for automatic inclusion in the table, but the author feels that these particular cases should be included for reference to the main text. The full profiles can be found at: http://www.biols.susx.ac.uk/Biochem/Woolfson/html/coiledcoils/cccat/mulal/stats/collated_tally_norm_rep.html

Furthermore, regarding the special residues Cys, Gly, and Pro, these occur with quite low frequencies in coiled-coil proteins (Conway and Parry, 1990, 1991; Lupas *et al.*, 1991; Parry, 1982; Woolfson and Alber, 1995) and tend only to be used to perform very specific roles in *de novo* coiled-coil designs. Gly and Pro are regarded as α-helical breakers and so are rarely chosen by designers for the central regions of coiled-coil structures. Cys can form disulphide bridges or be otherwise oxidized. It is usually avoided in design sequences unless it is specifically required for cross-linking polypeptide chains.

B. van der Waals' Forces, Steric Constraints, and Oligomer Specification

One property of hydrophobic interactions is that they tend not to be specific. In other words (and as evident in natural coiled-coil structures), the basic coiled coil pattern *HPPHPPP* is compatible with a number of helix-bundle quaternary structures. Dimeric, trimeric, tetrameric, pentameric, and dodecameric coiled coils are all known as are homomeric and heteromeric complexes and topologies with parallel, antiparallel, or mixed arrangements of helices (Burkhard *et al.*, 2001; Gruber and Lupas,

2003; Lupas, 1996; Walshaw and Woolfson, 2001a,b). Again, the reader is referred to Chapter 3 for a full description of coiled-coil structures.

Thus, the question in coiled-coil prediction and design is: what specific replacements are superimposed on the basic *HPPHPPP* pattern to direct the functional oligomerization state? This question was first tackled by Conway and Parry, who analyzed natural coiled-coil sequences that formed dimers and trimers (Conway and Parry, 1990, 1991). Woolfson and Alber (1995) advanced this approach by comparing amino-acid profiles for these two structures directly. The work that made the biggest impact on this issue, however, was the collaborative experimental study from the Kim and Alber laboratories using the GCN4 leucine-zipper peptide model system and mutants thereof.

The first experimental study was the high-resolution X-ray crystal structure of the leucine zipper peptide, GCN4-p1, itself (O'Shea *et al.*, 1991), which revealed that the packing geometries of residues at *a* and *d* were different. The significance of this finding for oligomer state selection became clearer with Harbury's studies of hydrophobic core mutants of the same peptide (Harbury *et al.*, 1993, 1994). As demonstrated by several protein engineering studies (Woolfson, 2001), hydrophobic residues within the cores of globular proteins tend to be very tolerant of substitution by other hydrophobic side chains. This is not the case for multichain coiled-coil structures. GCN4-p1, which forms dimers exclusively, has Leu at all four *d* positions, and 1 × Met, 1 × Asn, and 3 × Val at its *a* sites. Harbury has described multiple core mutants of GCN4-p1 in which all *a* residues (except Met-2) were simultaneously substituted for one of Ile, Leu, or Val, and all *d* residues were simultaneously substituted for one of same subset of aliphatic side chains (Harbury *et al.*, 1993). The peptides were generically named GCN4-p-*ad*. For example, GCN4-p-IL refers to the GCN4-p1 sequence with said *a* and *d* sites replaced by Ile and Leu, respectively. Seven of the nine possible peptides have been characterized. The comparison of peptides p-IL, p-II, and p-LI is interesting as they form dimers, trimers, and tetramers, respectively. Structures for the new trimeric (Harbury *et al.*, 1994) and tetrameric (Harbury *et al.*, 1993) forms are available and these have helped rationalize the different oligomer state selections on the basis of different packing arrangements made within the cores of the structures.

As O'Shea and colleagues (1991) have noted, the packing of side chains at the *a* and *d* sites of the GCN4-p1 dimer is different. At the *d* position, the $\alpha-\beta$ bond vector of the side chain points into the interface and directly towards the neighboring helix. This type of geometry, called perpendicular packing by Harbury, precludes β-branched residues such as Ile and Val from occupying these sites and favors Leu. This is the reason that the leucine zipper is a *leucine* zipper: the hallmark of the bZIP transcription

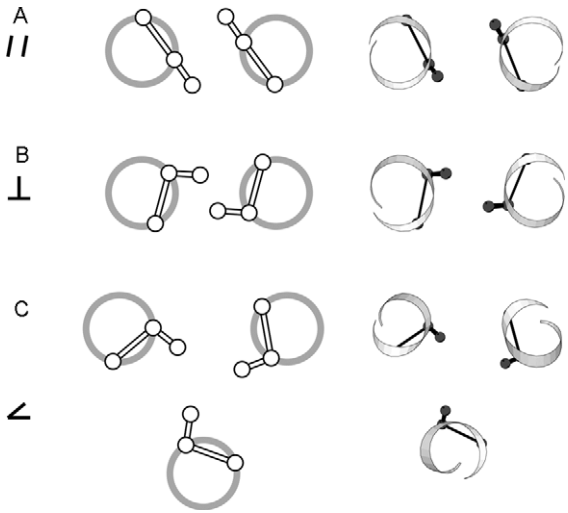


FIG. 2. Schematic diagrams (left) and experimental examples (right) for different core-packing geometries. (A) Parallel packing, showing an *a*-layer from the GCN4 leucine zipper (O'Shea *et al.*, 1991). (B) Perpendicular packing at a *d*-layer in the same structure. (C) Acute packing at a *d*-layer in a structure of a trimeric GCN4 mutant (Gonzalez *et al.*, 1996c). Reproduced from Walshaw and Woolfson (2001b).

factors is a run of four or so leucine residues spaced seven residues apart, and in a heptad repeat these fall at the *d* sites (Landschulz *et al.*, 1988). At the *a* sites, however, the $\alpha-\beta$ bond vector points out from the helical interface, and is termed *parallel packing*. As a result, the *a* site of dimeric coiled coils is much more permissive of amino acid substitution (Conway and Parry, 1990; Hu *et al.*, 1990; Woolfson and Alber, 1995). In hydrophobic side chains, the β-branched residues Ile and Val are favored at *a* because they contribute hydrocarbon back into the helical interface. These different geometries for the dimer case are shown in Fig. 2. In Harbury's p-LI tetramer, the packing geometries are reversed compared with the dimer. The packing at *a* is perpendicular, whereas the packing at *d* is parallel. This fits with the swap of Ile and Leu residues at *a* and *d* between the p-IL dimer and the p-LI tetramer.

In summary, the relative order of the β-branched residues and leucine at *a* and *d* has a major influence on oligomer state selection. The rule is that Leu prefers perpendicular packing whereas the β-branched residues, Ile and Val, prefer parallel packing (Table I). On this basis, it is of little surprise that a retro-GCN4-p1 sequence (i.e., with a = Leu and d = Ile) forms a tetramer and not a dimer (Mittl *et al.*, 2000).

So what happens in the trimer? GCN4-p-II crystallizes as a trimer (Harbury *et al.*, 1994). In this case, the packing at *a* and *d* is closely similar: it is intermediate between the perpendicular and parallel extremes (Fig. 2), and accordingly is referred to as acute packing. As a result, less selectivity for amino-acid type might be expected between the two sites. Indeed, this largely fits with Harbury's experimental studies and with analyses of natural trimeric coiled-coil sequences (Conway and Parry, 1991; Woolfson and Alber, 1995) (Table I). Taking these studies together, the rule of thumb of $a = d =$ Ile or Leu to specify trimers generally emerges. However, one further point is worth making here: the trimer state seems to be the default, and in the absence of strong sequence determinants of oligomer state, coiled-coil peptides and proteins will tend to form this structure. This is apparent from a number of studies using the GCN4-p1 background. First, Harbury's peptides with all $a =$ Val tend to form mixed oligomer states in solution (Harbury *et al.*, 1993). In addition, substitutions that replace the central *a* site, which is Asn in the wild-type sequence, form mixtures of oligomer states (Gonzalez *et al.*, 1996a,b,c). These results suggest that, at least for the *a* position, Val specifies the oligomer state very poorly compared with Ile. However, as described below, Val in combination with certain polar residues at *a*, as in wild-type GCN4-p1, does specify the dimer (Gonzalez *et al.*, 1996c; Lumb and Kim, 1995; Woolfson and Alber, 1995).

Walshaw and Woolfson (2001b) have analyzed the packing geometries of all coiled-coil structures identified in the Brookhaven Protein Data Bank (release #87), and have largely confirmed Harbury's experimental analysis. Furthermore, their work provided updated and structurally validated amino-acid profiles for the main coiled-coil quaternary structures. For instance, in parallel dimers the rule of $d =$ Leu combined with $a =$ Ile or Val, together with rules for polar residues at *a* (see below), largely holds up (Table I). These rules are, of course, rules of thumb and considerable variation is observed in natural sequences. Nonetheless, they are extremely powerful for specifying the oligomer state of designed coiled-coil peptides with confidence. Furthermore, several protein engineering and redesign studies have reproduced these observations and put measures on the thermodynamic effects of making certain hydrophobic mutations at the core *a* and *d* sites (Moitra *et al.*, 1997; Tripet *et al.*, 2000; Wagschal *et al.*, 1999). More recently, researchers have begun to use computational methods (Havranek and Harbury, 2003; Keating *et al.*, 2001) and to incorporate nonnatural amino acids to probe, understand, specify, and stabilize coiled-coil hydrophobic interfaces (Bilgicer *et al.*, 2001; Schnarr and Kennan, 2001, 2002, 2003; Yoder and Kumar, 2002).

Before leaving the subject of the hydrophobic component of coiled-coil interfaces, it is worth noting that hydrophobic residues at *e* and *g* and possibly at *b* and *c* also influence oligomer state selection. In coiled-coil sequences, the frequencies of hydrophobic residues at these sites increase with oligomer state (Conway and Parry, 1990, 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995). Likewise, in coiled-coil structures these sites are increasingly buried and involved in the interface (Harbury *et al.*, 1993; Kajava, 1996; Walshaw and Woolfson, 2003). Though these observations need further experimental testing and verification, potentially they provide a further route to designing coiled coils of different oligomer states.

C. *Salt-Bridge Interactions and Partner Selection*

It has long been appreciated that residues that flank the main hydrophobic seam of the coiled-coil interface (i.e., residues at *e* and *g* and, more latterly, those at *b* and *c*) are likely to contribute to coiled-coil specificity and stability (McLachlan and Stewart, 1975; McLachlan *et al.*, 1975). Specifically, in parallel coiled-coil structures, the *g* site of one helix is brought close to the *e* site of the successive heptad of a neighboring helix (Fig. 1C) in a so-called $g_n{:}e_{n+1}'$ interaction. As will be discussed later, in antiparallel structures the analogous interactions are $e{:}e'$ and $g{:}g'$ (Fig. 1D). This difference may be important in distinguishing between parallel and antiparallel structures in nature, and it also provides the protein designer with excellent positive and negative design rules.

At least from the analyses of amino-acid profiles for parallel dimers and trimers (Conway and Parry, 1990, 1991; Lupas *et al.*, 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995), it is beyond doubt that the placement of oppositely charged side chains at these sites must play a part in coiled-coil assembly, structure, and stability. Likewise, the proximity of oppositely charged side chains emanating from these sites in experimentally determined structures and the postulated resulting salt bridges add further compelling evidence for the importance of such interactions (O'Shea *et al.*, 1991). Nonetheless, and despite many experimental studies to probe these interactions, the issue remains controversial (Lavigne *et al.*, 1996; Lumb and Kim, 1996). Specifically, the question is whether $g_n{:}e_{n+1}'$ interactions stabilize and specify coiled-coil structures, or if they simply help specify partner selection and helix orientation.

Although this may seem like dodging an important issue—clearly in protein design it is important to understand how protein sequence relates to both protein structure *and* stability—for the purposes of this review, it is an academic question. This is for two reasons, which will become clear in

the examples of successful coiled-coil designs. First, stability tends not to be an issue when designing coiled coils, as plenty can be gained from the core interactions alone. Second, through the elegant studies of O'Shea and colleagues on the Fos:Jun system, it is clear that charge-charge complementarity in coiled-coil interfaces plays a significant role in partner selection (O'Shea *et al.*, 1989, 1992).

On this basis, coiled-coil designers often place oppositely charged residues Lys and Glu at complementary *g* and *e* sites to direct the assemblies that they are targeting (positive design), and make similar potential interactions in alternative structures repulsive (negative design).

D. Buried Polar Groups: The Icing on the Cake

Returning to the *a* and *d* sites of coiled-coil interfaces, they are not the exclusive province of hydrophobic residues: polar residues are found here and, in many cases, they are highly conserved. In retrospect, this is also apparent in the amino-acid profiles of dimeric and trimeric coiled coils (Conway and Parry, 1990, 1991; Lupas *et al.*, 1991). Furthermore, inspection of these profiles reveals trends in the data; for instance, basic residues occur frequently at the *a* sites of dimeric coiled coils (Conway and Parry, 1990). The real importance of such inclusions, however, only became apparent with the determination of the structure of the leucine-zipper peptide GCN4-p1 (O'Shea *et al.*, 1991).

Position 16 of GCN4-p1—an *a* site—is Asn, which is conserved across many leucine-zipper sequences. In the structure, this residue is buried and makes an asymmetric side chain-side chain hydrogen bond with the corresponding Asn in the partner strand. This Asn-Asn′ pair appears to play a number of roles, some of which have been subsequently verified experimentally. First, the inclusion of the pair in the otherwise hydrophobic core is destabilizing, and replacement of the residue by the canonical Val increases coiled-coil stability considerably (Harbury *et al.*, 1993). There may be biological significance to this in, for instance, modulating dimer stability for control and protein turnover in the cell. However, what is more dramatically obvious in the Asn16Val mutant is its loss of oligomer-state specificity: it forms mixtures of oligomers (Harbury *et al.*, 1993). Thus, second, Asn-16 is essential for specifying the dimer state of GCN4-p1. Third, the residue may be important in destabilizing other alternative structures, such as antiparallel and out-of-register structures, both of which would result in two buried and noncomplemented Asn side chains.

The importance of Asn at *a* in specifying the dimer is further demonstrated by the comparison of amino-acid profiles for dimeric and trimeric coiled coils (Woolfson and Alber, 1995). In the normalized profiles, Asn at

a is significant in dimers, but not in trimers. Similarly, and following on from the earlier studies of Conway and Parry (1990), the profiles from Woolfson and Alber show that the basic residues Lys and Arg occur significantly at the *a* sites of dimers, but not in trimers. With regard to promoting the trimer the residues Gln, Ser, and Thr are all favored at the *a* sites of trimer sequences, but not in dimers. A number of experimental studies lend support to these correlations, and provide structural rationales for the inclusion and accommodation of polar residues in coiled-coil interfaces (Akey *et al.*, 2001; Campbell and Lumb, 2002; Campbell *et al.*, 2002; Eckert *et al.*, 1998; Gonzalez *et al.*, 1996c; Lumb and Kim, 1995; Oakley and Kim, 1998). For instance, inclusion of Lys at *a* appears to specify dimers, not because of specific interhelix interactions in this state, but because the charge side chain cannot escape the core of higher-order states such as the trimer, which are therefore compromised (Gonzalez *et al.*, 1996c).

In summary, the inclusion of the polar residues, even when hydrogen-bonding potentials are satisfied, is destabilizing. However, though the burial of a polar residue may destabilize the target structure, it may destabilize alternative structures even more (Gonzalez *et al.*, 1996c). In this respect, the use of buried residues provides a very powerful negative-design principle for coiled-coil design.

III. Key Coiled-Coil Designs

The author recognizes that some readers will find it difficult to visualize the sequences given below in the context of coiled-coil structures. For this reason, it is suggested that readers consult the original articles, nearly all of which give annotated helical-wheel diagrams for the design sequences or, preferably, construct their own annotated helical wheels using the sequences given and blank diagrams of the type shown in Fig. 1C and D.

A. Parallel Structures

1. The Original Hodges Design: An 86-residue Analog of Tropomyosin (1981)

To the author's knowledge, Hodges and colleagues presented the first discrete coiled-coil peptide of *de novo* design (Hodges *et al.*, 1981). This design followed earlier studies by the group on polymers of the heptapeptide repeat sequence KLESLES. The design principles, which are based on an analysis of the amino-acid composition of tropomyosin, are clear and were well ahead of their time. The peptide comprised three blocks,

A (KCAELEG), B (KLEALEG), and C (KLEALEGK), assembled in the order AB_4C to give a 43 residue peptide, which was air-oxidized through the N-terminal Cys to form an 86-residue covalent dimer. The main features of this design were the straightforward $a = d =$ Leu core, the complementary *g*:*e′* Lys:Glu pairs, and the unusual Gly residues at *f*, which were included for synthetic reasons. Note that the heptad repeat is denoted *gabcdef* in this design. The dimer exhibited considerable helicity and thermal stability consistent with a fully folded coiled coil. This original design has spawned considerable work in coiled-coil design and redesign from the Hodges group (Kohn and Hodges, 1998) and others (Schneider *et al.*, 1998).

2. *Coil-Ser: A Peptide Designed as Parallel Homodimer that Forms an Up-Up-Down Homotrimer (1990)*

O'Neil and DeGrado (1990) adapted the original Hodges sequence to engineer a coiled-coil-based host-guest system for determining a thermodynamic scale of the helix-forming tendencies of the amino acids. Based on Hodges' building block—KLEALEG—they described a four-heptad peptide. Some significant changes were made, however. The first *a* position and the last *f* site were Trp and His, respectively, to facilitate spectroscopic characterization. Also, the glycine residues at the *f* sites of first and third heptad were replaced by Lys to improve helical propensity, and the remaining *f* position of the second heptad was used as the guest site and substituted by all 20 amino acids plus Aib. The stabilities of the resulting 21 peptides were determined by urea denaturation experiments. These showed a striking correlation with statistical measures of α-helix propensity.

The determination of the crystal structure of coil-Ser provided an interesting twist on this story (Lovejoy *et al.*, 1993). Rather than the parallel dimer expected, coil-Ser formed an up-up-down homotrimer (Fig. 3). Trimerization in solution was also confirmed by analytical ultracentrifugation. The likely reasons for this were the predominantly all-Leu core, and the inclusion of Trp at the first *a* position. As noted above, the $a = d$ core favors trimers, and Trp is too bulky for all three Trp residues to be accommodated within one coiled-coil layer, hence the switch of one helix to give the up-up-down arrangement. As noted by DeGrado and colleagues, this result did not affect the scale of helix-forming tendencies originally reported using the coil-Ser system (Lovejoy *et al.*, 1993).

3. *Peptide Velcro: A Parallel Heterodimer (1993)*

The ''Peptide Velcro'' design by O'Shea *et al.* (1993) was based on their experience with the GCN4-p1 and Fos/Jun systems. The core design—which preceded Harbury's work on oligomer-state selection (Harbury *et al.*,

FIG. 3. The ribbon structure of coil-Ser [pdb identifier, 1cos; (Lovejoy *et al.*, 1993)]. The structure reveals a mixed parallel/antiparallel three-helix bundle, which is influenced by the three core tryptophan side chains, shown in space-filling representation.

1993)—was influenced by the foregoing Hodges and DeGrado designs and was straightforward, with four $a = d =$ Leu heptads. However, the system did not form trimers because of the GCN4-inspired inclusion of Asn at the central *a* position (O'Shea *et al.*, 1991). In one partner, Acid-p1, Glu were placed at all *e* and *g* sites while its complement, Base-p1, had Lys at these positions. Thus, the idea of the design was to draw the two peptides together with a full set of complementary $g_n{:}e_{n+1}'$ pairings between the partner peptides. Consistent with the design principles, the isolated peptides did not fold in water but, when mixed, formed a stable 1:1 complex that is fully helical and unfolded cooperatively. The dissociation constant for this design is an impressive 30 nM (20 °C, pH 7). Additional experiments revealed that, as in the Fos/Jun system, repulsive electrostatic interactions in the alternative homodimers direct heterospecificity more than the favorable electrostatic interactions in the targeted heterodimer.

Interestingly, when the buried Asn was replaced by Leu, structural specificity was lost and the peptides formed a heterotetramer without fixed helix-helix orientations (Lumb and Kim, 1995). In a further report on this system, Sia and Kim (2001) showed that the two copies of Acid-p1 in the heterotetramer could be replaced by a peptide, D-Acid, in which all of the residues were made from D-amino acids. This clever redesign was guided by helical-net diagrams that considered core packing in the D/L structure. These revealed *a:d′* rather than *a:a′* and *d:d′* layers in the core, which were accommodated in the new design by a half-heptad shift of a redesigned L-Base sequence.

While on the topic of heterodimer design, Vinson and coworkers have had considerable success in making heterodimerizing leucine zippers with an impressive range of properties, including very tight binding (Moll *et al.*, 2001; Vinson *et al.*, 1993). Hodges and colleagues have presented a series of related designs, E/K coils, which they propose for use in biotechnology as, for example, capture reagents for affinity chromatography and biosensor-based applications (Chao *et al.*, 1998; Litowski and Hodges, 2001, 2002).

4. ABC: A Parallel Heterotrimer (1995)

The ABC design of Nautiyal *et al.* (1995) represented a step up in complexity and a considerable increase in the degree of difficulty over foregoing heterodimeric coiled-coil designs. Again, this was a four-heptad design with a straightforward $a = d =$ Ile core and a single polar residue (Gln) at *a* as a failsafe for parallel trimer formation. With only two charges to play with and three helical interfaces to specify, the selection of *e* and *g* residues used an algorithm to choose from more than 16 million peptide combinations. The solution maximized salt bridges in the targeted parallel heterotrimer, and minimized them in all alternative combinations. In this respect, the approach used both positive and negative-design principles. Two solutions to the problem were returned: the one synthesized, and its mirror image reversed the order of helical packing. One failing of the design was that the individual peptides did trimerizes, as did the binary combinations. However, none of these combinations matched the heterotrimer in thermal stability. An important and impressive aspect of this design was that its crystal structure confirmed all features of the design, including the correct A–B–C helical packing arrangement (Nautiyal and Alber, 1999).

5. A Heterotetramer: Extending the Coiled-Coil Interface (1996)

Based on the natural sequence of the core domain of the Lac repressor, which tetramerizes, Fairman *et al.* (1996) have engineered two 21-residue peptides, Lac21E and Lac21K. In isolation, the peptides did not fold

appreciably, but when combined they formed thermostable tetramers. In these short sequences, the *d* residues were Leu, and the three *a* sites were Met, Leu, and Val, respectively. Interestingly, the *e* and *g* positions were combinations of Ala, Glu, Gly, and Ser. This left specifying the structure down to charged residues at the *b* and *c* positions, which were all Glu in Lac21E and all Lys in Lac21K. This idea of extending the coiled-coil interfaces beyond the *a* and *d* seam and the *e* and *g* flanking residues to *b* and *c* has been noted elsewhere for coiled-coil structures (Harbury *et al.*, 1993; Kajava, 1996; Walshaw and Woolfson, 2003) and sequences (Conway and Parry, 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995).

6. Anti-APCp1: A Coiled-Coil Probe for the APC Tumor Suppressor Protein (1998)

Mutations of the human APC gene are associated with both sporadic and familial forms of colon cancer. The APC protein is a large, multidomain protein that has a 55-residue, N-terminal dimeric coiled coil (APC-55). Alber and colleagues used rules of thumb and those derived from an analysis of the covariation of *a:a′* and *d:d′* pairs in the cytokeratins (which form obligate heterodimers) to create a mutant of APC-55, anti-APCp1, as a potential probe for the APC protein (Sharma *et al.*, 1998).

In all, 20 mutations were described that can be grouped as follows. First, five changes were made at the *a* and *d* sites based on the analysis of the keratins. Second, eight changes were made at the *e* and *g* sites to introduce salt bridges to favor the anti-APCp1:APC-55 heterodimer and to destabilize homodimerization of anti-APCp1. Finally, seven changes were made at the *e*, *f*, and *g* positions to improve the charge and the helicity of the peptide and to introduce chromophores. Although anti-APCp1 did form stable oligomers, these were higher orders than the dimer. Moreover, when mixed with APC-55, a stabilized heterodimer was formed as judged by solution-phase biophysics and gel electrophoresis. In addition, the peptide probe pulled down wild-type and mutant forms of APC from extracts of cancer cell lines.

7. RH: Designed Right-Handed Coiled Coils (1998)

Harbury *et al.* (1998) described a series of peptides designed to form dimeric, trimeric, and tetrameric helical bundles with right-handed supercoils. To achieve right-handed coiled coils, rather than canonical left-handed structures, an *HP* pattern based on an 11-residue *abcdefghij* repeat was used as a template. Combinations of the hydrophobic residues Ala, Ile, Leu, Val, allo-Ile, and nor-Val were considered for the *a*, *d*, and *h* sites. The

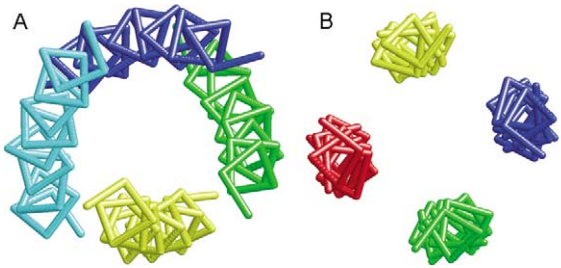


FIG. 4. (A) An end view of the GCN4-pLI (Harbury *et al.*, 1993) tetramer backbone showing the left-handed supercoil. (B) Similar projection for RH4 (Harbury *et al.*, 1998) showing the slighter supercoil in the opposite sense.

novel inclusion of the nonstandard residues was based on preliminary modeling studies. A computer algorithm was written to choose optimal combinations of these residues for the targeted dimer, trimer, and tetramer states. An additional novel feature of the design protocol, which advanced protein design in general, was the inclusion of some backbone flexibility in the modeling algorithm. Together, these novel features allowed the sequence of *a*, *d*, and *h* sites to be optimized for the dimer, trimer, and tetramer states. Solution phase characterization of three corresponding peptides—RH2, RH3, and RH4—confirmed the designs as stable, cooperatively folded helical bundles of the correct oligomer state, though RH2 was only marginally stable. Most importantly for the confirmation of the design process, the structure of RH4 was described. This revealed a right-handed coiled-coil structure with core packing and other features that matched the design model in atomic detail. The structure of RH4 is compared with the left-handed coiled-coil tetramer GCN4-p-LI in Fig. 4.

On the theme of noncanonical coiled coils, Hicks *et al.* (2002) have investigated the effects of a range of inserts—of between one and seven Ala residues—at the center of an otherwise canonical designed four-heptad leucine-zipper system. All of the inserts were destabilizing, but the four-residue insert, which results in a heptad-hendecad-heptad-heptad (7-11-7-7) sequence, was the least destabilizing. Similarly, though for different reasons, Hodges and colleagues have also made and characterized canonical coiled coils as hosts for peptide inserts (Kwok *et al.*, 2002).

8. SAF: An Offset Heterodimer that Promotes Fiber Formation (2000)

Pandya *et al.* (2000) have described the design of a sticky-ended, or offset, leucine-zipper-based heterodimer (Fig. 5A). This is unusual because all known natural coiled coils are blunt ended, presumably as a

consequence of the hydrophobic effect, although Talbot and Hodges (1982) have commented on possible (though unintentional) promiscuous sticky-ended assembly in early coiled-coil designs. In the design from Pandya and colleagues, the free ends of the sticky-ended dimer were made complementary to promote longitudinal assembly into extended self-assembled protein fibers (SAFs). The features of the design included a core that promotes dimerization with a = Ile and d = Leu, residues at e and g that promote sticky-end assembly, N-terminal halves of the peptides with $e = g =$ Lys, and corresponding residues in the C-terminal halves that were Glu. However, the key novel feature of this design is the offset placing of complementary Asn residues at a C-terminal a site in one peptide and an N-terminal a site in the other. In the absence of a high-resolution structure, the design was confirmed using a combination of spectroscopy (circular and linear dichroism), electron microscopy, and X-ray fiber diffraction (Pandya *et al.*, 2000). The peptides assembled into linear, nonbranched fibers tens of microns long and approximately 50 nm thick (Fig. 5B). The thickness of the fibers was unexpected—coiled-coil dimers are typically ≈2 nm thick—and reflected assembly of 2 nm protofibrils into the matured fibers, which in turn indicated that the protofibrils with helically repeated building blocks were inherently sticky.

The SAF design has been built upon considerably over the past two years (Ryadnov and Woolfson, 2003a,b, 2004). Others are pursuing coiled-coil designs as a route to self-assembling fibrous structures. Indeed, the first example dates back to 1997 (Kojima *et al.*, 1997). More recently, Kajava and colleagues have described the design and characterization of self-assembling fibers based on a novel pentameric coiled-coil building block (Kajava *et al.*, 2004; Melnik *et al.*, 2003; Potekhin *et al.*, 2001), and Conticello and colleagues have described fibrous structures based on a

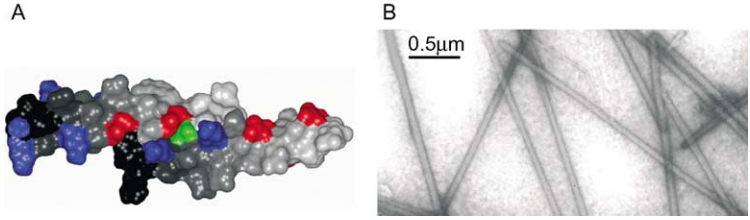


FIG. 5. (A) Molecular model for the designed SAF sticky ended heterodimer. The Lys and Glu residues at the e and g sites are blue and red, respectively; the buried, offset Asn residues at a are green. Adapted from Pandya *et al.* (2000). (B) Negative-stain transmission electron micrograph image for fibers assembled from the sticky ended building blocks. Adapted from MacPhee and Woolfson (2004).

single, 42-residue leucine zipper-based helix (Zimenkov *et al.*, 2004). On a related theme, Ghosh and colleagues have reported fibers produced by combining basic (EZ) and acidic (KZ) $a = d =$ Leu coiled-coil peptides (Zhou *et al.*, 2004). The novelty in this work was that each peptide was tethered to a core dendrimer (D) to give constructs D-EZ_4 and D-EK_4, which were then used in assembly. Along with other work—for example, in the development of coiled coil-based hydrogel systems (Petka *et al.*, 1998; Wang *et al.*, 1999)—these papers represent an exciting beginning for the application of coiled-coil design in the area of the rational design and exploitation of nanostructured biomaterials. Some of these papers and this field have been reviewed recently (MacPhee and Woolfson, 2004; Yeates and Padilla, 2002).

9. The Shortest Designed Coiled-Coil Peptide (2000)

In three recent publications (Burkhard *et al.*, 2000b, 2002; Meier *et al.*, 2002), Burkhard and colleagues have described the design, redesign, and characterization of a two-heptad coiled-coil system, Suc-DELEARIRELE ARIK-NH_2. This unit was stabilized by hydrophobic interactions ($a =$ Ile, $d =$ Leu) and a network of salt bridge interactions. Initial characterization of the peptide showed that it was helical and dimeric. However, at increased ionic strength, the oligomer state switched to trimer. A structure for this form is available. Encouragingly, the introduction of improved saltbridge interactions, Suc-EELRRRIEELERRIR-NH_2, did further specify the dimer state in solution, though the structure deposited in the Protein Data Bank for this peptide is still for the trimer. Likewise, removal of salt bridges, Suc-DELERAIRELAARIK-NH_2, resulted in the loss of coiled-coil specificity and the formation of a noncoiled-coil octamer. It is probable that the shortness of this design contributes to the change in oligomer state and overrides the influence core residues, ($a =$ Ile, $d =$ Leu), which would normally specify dimer. In other words the relative stabilities of alternative structural states of the unit is probably marginal and, thus, small changes in the noncovalent interactions affect the equilibria between these states significantly.

10. Coiled Coils Designed to Switch Conformational State (2002 and 2003)

Coiled-coil motifs have been known to play roles in conformational switching in natural proteins for some time (Oas and Endow, 1994). The key examples are influenza hemagglutinin (Bullough *et al.*, 1994; Carr and Kim, 1993; Carr *et al.*, 1997), and the heat shock transcription factor (Rabindran *et al.*, 1993). Furthermore, an engineered form of GCN4-p1, with Asn-16 replaced by Ala, switches from dimer to trimer upon addition of

cyclohexane or benzene (Gonzalez *et al.*, 1996b). In the presence of benzene, the mutant peptide crystallizes as a parallel trimer with a single benzene molecule bound at a cavity bordered by Leu-12, Ala-16, and Leu-19.

Ciani *et al.* (2002) have described the *de novo* design of Template-α, a 29-residue peptide with a canonical dimeric leucine zipper repeat, *a*AALEQK (where $a =$ Val, Lys and Ile). As expected, the peptide folded to a helical dimer that unfolded reversibly and in a concentration-dependent manner upon heating in solution. The peptide design was also compatible with an antiparallel β-hairpin structure. To further enhance β-sheet propensity, a mutant peptide, Template-αT (in which the Gln residues at the *f* sites of Template-α were replaced by Thr) was made. Consistent with this dual-sequence design principle, thermal unfolding of the mutant resulted in the formation of β-structure as judged by CD spectroscopy, and amyloid-like structures as observed by transmission electron microscopy. On a similar note, Kammerer *et al.* (2004) have described a shorter designed coiled-coil peptide, *cc*β, which crystallized as a trimer and transformed to amyloid-like structures upon heating. These systems present possibilities for assessing how α-helical structures switch to β-sheet-based amyloid-like structures, and this may shed light on how natural proteins transform to amyloid in certain diseases.

Related to these ideas, Pandya *et al.* (2004) have described the design of an antiparallel coiled-coil (helix-loop-helix) peptide, which is stabilized by a disulfide bridge between the termini of the peptide. Reduction of the disulfide triggered a switch to a dimeric leucine zipper.

11. Belt-and-Braces: A Leucine Zipper-Based Nanoscale Linker System (2003)

The design and engineering of systems to control the assembly of structures and functional modules with nanometer precision is one goal of the burgeoning discipline of nanotechnology. Designed coiled-coil systems offer great possibilities in this area for several reasons. First, coiled coils can be considered as relatively stiff nanoscale rods. Second, there is a precise relationship between peptide length and coiled-coil length, as each heptad meters out $\approx$1 nm. Finally, coiled coils fold and self-assemble to stable structures at μM concentrations, or lower in water.

Ryadnov *et al.* (2003) have designed a coiled coil-based nanoscale linker system dubbed Belt-and-Braces. The system was novel in a number of respects. Though based on a leucine-zipper dimer design, it was a ternary system in which one peptide (the ''Belt'') templated the assembly of two half-sized peptides (the ''Braces''); thus, the system was the first and simplest example of a coiled-coil vernier assembly (Kelly *et al.*, 1998). The Belt-and-Braces design employed all of the key design rules for a

parallel coiled-coil dimer: that is, a = Ile, d = Leu, and the g_n:e_{n+1}' pairs are Lys:Glu or Glu:Lys. Ternary complexation was directed by the Belt, which was a six-heptad acid peptide with the general (g–f) heptad repeat, EIAALEQ, that templated the two three-heptad Braces B_N and B_C, with the repeat sequence KIAALKQ. Specificity was achieved by complementary Asn residues at the fifth a site of the Belt and the central a site of B_C. Finally, the N- and C-termini of B_N and B_C, respectively, were extended by Cys-Gly-Gly- and -Gly-Gly-Cys units. The construct was expected to span 6 to 7 nm. The folding and assembly of the naked Belt-and-Braces peptides was confirmed by a combination of CD spectroscopy, analytical ultracentrifugation (AUC), and surface plasmon resonance (SPR) using Biacore. Moreover, the assembly of bound cargo was demonstrated as follows: the Cys-termini of the Brace peptides were labeled with 15 nm colloidal gold particles and mixed with the Belt. Transmission electron microscopy (TEM) visualization of the resulting assemblies revealed networks of nanoparticles separated by approximately 7 nm consistent with the design.

Stevens *et al.* (2004) reported a similar leucine-zipper-like linker system. This comprised two six-heptad peptides (d = Leu, a = Val, Ala, and Ile), one of which was basic with predominantly Lys at e and g, while the other was acid with predominantly Glu at e and g. Assembly was confirmed by CD spectroscopy and visualized by coupling the peptides to gold nanoparticles followed by TEM. A nice touch in this work was the use of different size nanoparticles to create binary assemblies including satellite structures, in which 8.5 nm particles (derivatized with the acid peptide) were organized around 53 nm particles (derivatized with the basic peptide).

The designs by Hodges and colleagues in the section on heterodimeric coiled-coil systems (Chao *et al.*, 1998; Litowski and Hodges, 2001, 2002), and by Ghosh *et al.* (2000) in the next section, provide other examples of coiled-coil systems as potential peptide linkers.

B. *Antiparallel Structures*

With the exception of the discussion centered on the crystal structure of coil-Ser, which forms a mixed parallel/antiparallel trimer (Lovejoy *et al.*, 1993), this review has so far ignored antiparallel coiled-coil structures. This is for a number of reasons. First, our understanding of antiparallel coiled coils is not as advanced as that for the parallel structures (Oakley and Hollenbeck, 2001; Walshaw and Woolfson, 2001b). Second, and as a result of the first point, fewer design attempts have targeted antiparallel coiled coils. Finally, there is a very good recent review on antiparallel coiled-coil structures and design (Oakley and Hollenbeck, 2001). Nonetheless, this Chapter would not be complete without the inclusion of these design

attempts. Beforehand, however, the main differences between antiparallel and parallel coiled-coil structures are summarized. Because of the relative dearth of experimental data on other antiparallel structures, discussion is limited to heptad-based antiparallel dimers that, like the parallel structures, have left-handed supercoils.

In order to maintain the register of hydrophobic residues between two antiparallel coiled-coil strands, *a* residues from one strand must pair with *d* residues of the other (Fig. 1). Thus, whereas in parallel dimers alternating core layers comprise *a:a′* and *d:d′* residues, antiparallel dimers have alternating *a:d′* and *d:a′* layers. On this basis, the stereochemical packing of mixed *a:d* layers might be expected to differ from those in the symmetric layers of parallel structures. However, analysis of the known antiparallel dimer structures show that packing geometries of the *a* and *d* side chains can still be considered as roughly parallel and perpendicular, respectively (Walshaw and Woolfson, 2001b), although both distributions of angles are shifted slightly towards acute (trimer) packing. This is reflected in the amino-acid profile for antiparallel dimers (Walshaw and Woolfson, 2001b) (Table I). Briefly, this profile is a somewhat dampened version of the parallel dimer profile, with a preference for Leu at *d* and Leu/Ile at *a*. The interchain interactions between the core-flanking *e* and *g* residues are, however, very different: in parallel structures g_n and e'_{n+1} interactions are made across the interface, whereas in antiparallel structures it is *g:g′ and e:e′*. This provides the protein designer with a negative-design rule and a route to designing antiparallel coiled coils, avoiding alternative parallel structures.

1. *The First Hodges Construct (1993)*

As with the first design for a parallel coiled coil, Hodges and colleagues were also the first to describe a construct for an antiparallel coiled coil (Monera *et al.*, 1993). This was based on the (*g–f*) heptad repeat KLEALEG (Hodges *et al.*, 1981). To construct the antiparallel system, several changes were made to the design. First, Leu-16, an *a* site, was replaced by Ala. Second, two peptides were used, incorporating Cys at the first *a* site (peptide C2A16) and the last *d* site (peptide C33A16). Both peptides were five heptads long. The two peptides were mixed under denaturing conditions and air-oxidized to give a 1:1:2 mixture of the two forced-parallel homodimers and the forced-antiparallel heterodimer. Note that this construct did not consider electrostatic pairings between the *e* and *g* sites: the *g:g′* and *e:e′* pairs in the antiparallel construct are all repulsive, whereas the *g:e′* pairs in the parallel forms are attractive. Characterization of the forced-parallel and antiparallel constructs reflected this. With respect to heat and urea-induced denaturation, the forced-antiparallel dimer was less stable than either of the covalent parallel dimers; although the relative

stabilities with respect to guanidinium hydrochloride-induced denaturation were reversed, this is explained by a salt effect. Furthermore, when the covalent dimers were made under benign (folding) conditions, only the parallel dimers were formed; equilibration of the forced-antiparallel dimer with disulfide-exchange reagents returned the parallel dimers. Nonetheless, the antiparallel construct made was helical, cooperatively folded, and represented the first engineered antiparallel coiled coil. Furthermore, in the same paper, Hodges and colleagues recognized the error in their design and described a final construct in which charged pairs were optimized to favor the antiparallel heterodimer. They used these constructs in several subsequent papers to probe sequence-to-structure/stability relationships in parallel and antiparallel coiled coils (Monera *et al.*, 1994, 1996a,b).

2. *CCSL: A Coiled-Coil Stem Loop Structure of de novo Design (1994)*

Myszka and Chaiken (1994) described the rational design of a single-chain antiparallel coiled-coil structure. This design was very well conceived. The peptide was 56-residues long and it had two 25-residue coiled-coil sequences separated by a flexible and functional (RGD-containing) six-residue loop. Except for the first *a:d′* layer (which is Cys:Cys), the *d:d′* and *a:d′* layers were Leu:Leu and Val:Val, respectively. Furthermore, the *e* and *g* sites in the N-terminal half were all Glu, whereas those in the C-terminal half were Lys to facilitate favorable *g:g′* and *e:e′* interactions. The unit was stapled by a disulfide bond between Cys-1 and Cys-56 at *a* and *d′* sites, respectively, and the remaining *b, c*, and *f* sites were Ala and Ser. CCSL was soluble, monomeric, approximately 80% helical, and unfolded cooperatively. Finally, the peptide was active, competing with fibrinogen for the GPIIbIIIa receptor.

Based on very similar principles, though more recently, Suzuki and Fujii (1999) designed a helix-loop-helix peptide. On a related theme, the aforementioned design from Pandya *et al.* (2004) for a helix-loop-helix peptide stabilized by a disulfide bridge used similar ideas, although the final sequence was very different from the Myszka and Chaiken and the Suzuki and Fujii peptides, as it was also made compatible with a parallel coiled-coil dimer to promote conformational switching.

3. *An Antiparallel Variant of Peptide Velcro (1998)*

Oakley and Kim (1998) have described a simple and elegant experiment to switch the aforementioned Peptide Velcro from a parallel heterodimer to an antiparallel structure. The strategy was straightforward and resulted in a powerful design rule for specifying antiparallel structures. The Acid-p1

and Base-p1 sequences of Peptide Velcro were redesigned so as to mismatch the Asn residues at their central *a* sites; in Acid-a1 Asn is placed at the third *a* site, while in Base-a1 it is at the second *d* site. In the targeted antiparallel structure, these sites form an *a:d′* layer. The solution phase characterization indicated that the favored complex was the antiparallel heterodimer. This study added to the rules for coiled-coil assembly and specification because an Asn:Asn pair at *a:d′* was unprecedented in natural coiled-coil structures (Oakley and Hollenbeck, 2001). Oakley and colleagues have extended this work to improve the designs of Acid-a1 and Base-a1, using both positive and negative design principles to improve the *g:g′* interactions to favor the antiparallel orientation over the competing parallel arrangement (McClain *et al.*, 2001).

4. Antiparallel Leucine-Zipper-Directed Protein Fragment Complementation (2000)

A number of proteins are known that can be split into two fragments, which when recombined form functional, though usually less stable, folded structures (Michnick, 2001). This is known as *fragment complementation*. For proteins that cannot be split and reconstituted so straightforwardly, there is the possibility that reassembly might be directed by additional reagents such as ligands. Such systems present a potential route to new biosensors.

Ghosh *et al.* (2000) have described an elegant system in which two fragments of the green fluorescent protein (GFP) from *Aequorea victoria* could be recombined only in the presence of helical tags engineered at the new termini created by the split. The helices were two halves, NZ and CZ, of a designed antiparallel coiled-coil dimer. The peptides were 29 and 30 residues long respectively. The cores were all Leu, except the second *d* site of NZ and the third *a* of CZ, which were Asn. All of the potential *e:e′* and *g:g′* pairs were Lys:Glu, and the remaining *b*, *c*, and *f* sites were combinations of Ala, Gln, and Lys with a single Trp chromophore at an *f* site in each peptide. Reconstitution of GFP was demonstrated both *in vitro* using the purified protein fragments, and *in vivo* by co-expression of the two components in *E. coli*.

5. APH: A Designed Antiparallel Homodimer (2003)

The designs outlined above are effective all for heterodimers. The design of a homodimeric antiparallel coiled-coil is more challenging. For instance, specification via the burial of two Asn residues, one at *a* and the other at the corresponding *d′* site, is not possible. In a thoughtful design, Oakley and colleagues tackled this problem to generate APH (Gurnon *et al.*, 2003). APH is a 45-residue recombinant peptide

overexpressed in *E. coli*. The design had the following key features: although the core was predominantly Leu, there were two Ala:Ile pairs and two Leu:Arg pairs at *a:d′* layers. The argument for using the former pair was that the small Ala side chain helps accommodate Ile in the antiparallel form. The inclusion of Arg within the core followed work by the groups of Oakley and of Lumb indicating that this provided specificity at a smaller energetic price than buried Asn residues (Campbell *et al.*, 2002; McClain *et al.*, 2002). On the basis of amino-acid profiles (Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995) and experimental studies (Gonzalez *et al.*, 1996c), these inclusions probably also destabilize potential parallel dimers and trimers, respectively. Finally, to favor the antiparallel form over alternative parallel structures, all potential *e:e′* and *g:g′* interactions were made into attractive Glu:Lys pairs, whereas potential g_n:e_{n+1}′ interactions were made repulsive. The design held up to scrutiny by solution-phase biophysics.

IV. Summary

This review has attempted to build on the preceding Chapters that convey the richness of coiled-coil sequences and structures. With this backdrop of many potential coiled-coil structures and very many potential coiled-coil sequences, one aim of this Chapter has been to illustrate that sequence-to-structure relationships in coiled coils can be understood sufficiently to attempt with some confidence and, hence, with some degree of success, the rational design of coiled-coil structures. At present, this process is largely done on the backs of envelopes guided by good rules of thumb, intuition, and imagination although, and as noted, computer-aided designs are increasingly being reported. Nonetheless, the back-of-the-envelope approach has had many successes from designs for basic, naturally observed structures to more-imaginative ones that might be used as probes, tethers, molecule rulers, affinity matrices, and components of new materials. In these respects, coiled-coil design is arguably the most successful of all protein-design areas. As well as continuing such efforts, one challenge in the near future will be to formalize the design rules by linking them to thermodynamic measurements and employing them further in prediction and computer-aided design algorithms. The future of coiled-coil design is very exciting indeed.

Acknowledgments

The author thanks Andrei Lupas, Markus Gruber, David Parry, and members of my research group for their critical reading of and constructive comments on this manuscript.

References

Akey, D. L., Malashkevich, V. N., and Kim, P. S. (2001). Buried polar residues in coiled-coil interfaces. *Biochemistry* **40,** 6352–6360.

Beasley, J. R., and Hecht, M. H. (1997). Protein design: The choice of *de novo* sequences. *J. Biol. Chem.* **272,** 2031–2034.

Bilgicer, B., Fichera, A., and Kumar, K. (2001). A coiled coil with a fluorous core. *J. Am. Chem. Soc.* **123,** 4393–4399.

Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371,** 37–43.

Burkhard, P., Ivaninskii, S., and Lustig, A. (2002). Improving coiled-coil stability by optimizing ionic interactions. *J. Mol. Biol.* **318,** 901–910.

Burkhard, P., Kammerer, R. A., Steinmetz, M. O., Bourenkov, G. P., and Aebi, U. (2000a). The coiled-coil trigger site of the rod domain of cortexillin I unveils a distinct network of interhelical and intrahelical salt bridges. *Structure* **8,** 223–230.

Burkhard, P., Meier, M., and Lustig, A. (2000b). Design of a minimal protein oligomerization domain by a structural approach. *Prot. Sci.* **9,** 2294–2301.

Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001). Coiled coils: A highly versatile protein folding motif. *Trends Cell Biol.* **11,** 82–88.

Campbell, K. M., and Lumb, K. J. (2002). Complementation of buried lysine and surface polar residues in a designed heterodimeric coiled coil. *Biochemistry* **41,** 7169–7175.

Campbell, K. M., Sholders, A. J., and Lumb, K. J. (2002). Contribution of buried lysine residues to the oligomerization specificity and stability of the Fos coiled coil. *Biochemistry* **41,** 4866–4871.

Carr, C. M., Chaudhry, C., and Kim, P. S. (1997). Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc. Natl. Acad. Sci. USA* **94,** 14306–14313.

Carr, C. M., and Kim, P. S. (1993). A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73,** 823–832.

Chao, H. M., Bautista, D. L., Litowski, J., Irvin, R. T., and Hodges, R. S. (1998). Use of a heterodimeric coiled-coil system for biosensor application and affinity purification. *J. Chromatog. B Biomed. Life Sci.* **715,** 307–329.

Chothia, C., Levitt, M., and Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol.* **145,** 215–250.

Ciani, B., Hutchinson, E. G., Sessions, R. B., and Woolfson, D. N. (2002). A designed system for assessing how sequence affects alpha to beta conformational transitions in proteins. *J. Biol. Chem.* **277,** 10150–10155.

Cohen, C., and Parry, D. A. D. (1990). Alpha-helical coiled coils and bundles – how to design an alpha-helical protein. *Proteins* **7,** 1–15.

Cohen, C., and Parry, D. A. D. (1994). Alpha-helical coiled coils – more facts and better predictions. *Science* **263,** 488–489.

Conway, J. F., and Parry, D. A. D. (1990). Structural features in the heptad substructures and longer range repeats of two-stranded alpha-fibrous proteins. *Int. J. Biol. Macromol.* **12,** 328–334.

Conway, J. F., and Parry, D. A. D. (1991). Three-stranded alpha-fibrous proteins: The heptad repeat and its implications for structure. *Int. J. Biol. Macromol.* **13,** 14–16.

Dwyer, M. A., Looger, L. L., and Hellinga, H. W. (2004). Computational design of a biologically active enzyme. *Science* **304,** 1967–1971.

Eckert, D. M., Malashkevich, V. N., and Kim, P. S. (1998). Crystal structure of GCN4-pI(Q)I, a trimeric coiled coil with buried polar residues. *J. Mol. Biol.* **284,** 859–865.

Fairman, R., Chao, H. G., Lavoie, T. B., Villafranca, J. J., Matsueda, G. R., and Novotny, J. (1996). Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu(−)-Lys(+) ion pair interactions. *Biochemistry* **35,** 2824–2829.

Fujinaga, M., Berthetcolominas, C., Yaremchuk, A. D., Tukalo, M. A., and Cusack, S. (1993). Refined crystal-structure of the seryl-transfer rna-synthetase from thermus-thermophilus at 2-center-dot-5-angstrom resolution. *J. Mol. Biol.* **234,** 222–233.

Ghosh, I., Hamilton, A. D., and Regan, L. (2000). Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *J. Am. Chem. Soc.* **122,** 5658–5659.

Gonzalez, L., Brown, R. A., Richardson, D., and Alber, T. (1996a). Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. *Nature Struct. Biol.* **3,** 1002–1010.

Gonzalez, L., Plecs, J. J., and Alber, T. (1996b). An engineered allosteric switch in leucine-zipper oligomerization. *Nature Struct. Biol.* **3,** 510–515.

Gonzalez, L., Woolfson, D. N., and Alber, T. (1996c). Buried polar residues and structural specificity in the Gcn4 leucine-zipper. *Nature Struct. Biol.* **3,** 1011–1018.

Gruber, M., and Lupas, A. N. (2003). Historical review: Another 50th anniversary – new periodicities in coiled coils. *Trends Biochem. Sci.* **28,** 679–685.

Gurnon, D. G., Whitaker, J. A., and Oakley, M. G. (2003). Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc.* **125,** 7518–7519.

Harbury, P. B., Kim, P. S., and Alber, T. (1994). Crystal-structure of an isoleucine-zipper trimer. *Nature* **371,** 80–83.

Harbury, P. B., Plecs, J. J., Tidor, B., Alber, T., and Kim, P. S. (1998). High-resolution protein design with backbone freedom. *Science* **282,** 1462–1467.

Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between 2-stranded, 3-stranded and 4-stranded coiled coils in Gcn4 leucine-zipper mutants. *Science* **262,** 1401–1407.

Havranek, J. J., and Harbury, P. B. (2003). Automated design of specificity in molecular recognition. *Nature Struct. Biol.* **10,** 45–52.

Hellinga, H. W. (1997). Rational protein design: Combining theory and experiment. *Proc. Natl. Acad. Sci. USA* **94,** 10015–10017.

Hicks, M. R., Walshaw, J., and Woolfson, D. N. (2002). Investigating the tolerance of coiled-coil peptides to nonheptad sequence inserts. *J. Struct. Biol.* **137,** 73–81.

Hill, R. B., Raleigh, D. P., Lombardi, A., and Degrado, N. F. (2000). *De novo* design of helical bundles as models for understanding protein folding and function. *Acc. Chem. Res.* **33,** 745–754.

Hodges, R. S., Saund, A. K., Chong, P. C. S., Stpierre, S. A., and Reid, R. E. (1981). Synthetic model for 2-stranded alpha-helical coiled-coils – design, synthesis, and characterization of an 86-residue analog of tropomyosin. *J. Biol. Chem.* **256,** 1214–1224.

Hu, J. C., O'Shea, E. K., Kim, P. S., and Sauer, R. T. (1990). Sequence requirements for coiled-coils – analysis with lambda-repressor-Gcn4 leucine zipper fusions. *Science* **250,** 1400–1403.

Kajava, A. V. (1996). Modeling of a five-stranded coiled coil structure for the assembly domain of the cartilage oligomeric matrix protein. *Proteins* **24,** 218–226.

Kajava, A. V., Potekhin, S. A., Corradin, G., and Leapman, R. D. (2004). Organization of designed nanofibrils assembled from alpha-helical peptides as determined by electron microscopy. *J. Pept. Sci.* **10,** 291–297.

Kammerer, R. A., Kostrewa, D., Zurdo, J., Detken, A., Garcia-Echeverria, C., Green, J. D., Muller, S. A., Meier, B. H., Winkler, F. K., Dobson, C. M., and Steinmetz, M. O. (2004). Exploring amyloid formation by a *de novo* design. *Proc. Natl. Acad. Sci. USA* **101,** 4435–4440.

Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aebi, U., and Steinmetz, M. O. (1998). An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proc. Natl. Acad. Sci. USA* **95,** 13419–13424.

Keating, A. E., Malashkevich, V. N., Tidor, B., and Kim, P. S. (2001). Side-chain repacking calculations for predicting structures and stabilities of heterodimeric coiled coils. *Proc. Natl. Acad. Sci. USA* **98,** 14825–14830.

Kelly, T. R., Xie, R. L., Weinreb, C. K., and Bregant, T. (1998). A molecular vernier. *Tetrahedron Letts.* **39,** 3675–3678.

Kohn, W. D., and Hodges, R. S. (1998). *De novo* design of alpha-helical coiled coils and bundles: Models for the development of protein-design principles. *Trends Biotechnol.* **16,** 379–389.

Kojima, S., Kuriki, Y., Yoshida, T., Yazaki, K., and Miura, K. (1997). Fibril formation by an amphipathic alpha-helix-forming polypeptide produced by gene engineering. *Proc. Japan Acad. B Phys. Biol. Sci.* **73,** 7–11.

Kraulis, P. J. (1991). Molscript – a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24,** 946–950.

Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., and Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science* **302,** 1364–1368.

Kwok, S. C., Mant, C. T., and Hodges, R. S. (2002). Importance of secondary structural specificity determinants in protein folding: Insertion of a native beta-sheet sequence into an alpha-helical coiled-coil. *Prot. Sci.* **11,** 1519–1531.

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The Leucine Zipper – a hypothetical structure common to a new class of DNA-binding proteins. *Science* **240,** 1759–1764.

Lavigne, P., Sonnichsen, F. D., Kay, C. M., and Hodges, R. S. (1996). Interhelical salt bridges, coiled-coil stability, and specificity of dimerization. *Science* **271,** 1136–1137.

Lee, D. L., Lavigne, P., and Hodges, R. S. (2001). Are trigger sequences essential in the folding of two-stranded alpha-helical coiled-coils? *J. Mol. Biol.* **306,** 539–553.

Litowski, J. R., and Hodges, R. S. (2001). Designing heterodimeric two-stranded alpha-helical coiled-coils: The effect of chain length on protein folding, stability and specificity. *J. Pept. Res.* **58,** 477–492.

Litowski, J. R., and Hodges, R. S. (2002). Designing heterodimeric two-stranded alpha-helical coiled-coils – Effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity. *J. Biol. Chem.* **277,** 37272–37279.

Liu, J., Yong, W., Deng, Y., Kallenbach, N. R., and Lu, M. (2004b). Atomic structure of a tryptophan-zipper pentamer. *Proc. Natl. Acad. Sci. USA.* **101,** 16156–16161.

Liu, X. S., Fan, K., and Wang, W. (2004a). The number of protein folds and their distribution over families in nature. *Proteins* **54,** 491–499.

Looger, L. L., Dwyer, M. A., Smith, J. J., and Hellinga, H. W. (2003). Computational design of receptor and sensor proteins with novel functions. *Nature* **423,** 185–190.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., Degrado, W. F., and Eisenberg, D. (1993). Crystal-structure of a synthetic triple-stranded alpha-helical bundle. *Science* **259,** 1288–1293.

Lumb, K. J., Carr, C. M., and Kim, P. S. (1994). Subdomain folding of the coiled-coil leucine-zipper from the bZIP transcriptional activator GCN4. *Biochemistry.* **33,** 7361–7367.

Lumb, K. J., and Kim, P. S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled-coil. *Biochemistry* **34,** 8642–8648.

Lumb, K. J., and Kim, P. S. (1996). Interhelical salt bridges, coiled-coil stability, and specificity of dimerization – Response. *Science* **271,** 1137–1138.

Lupas, A. (1996). Coiled coils: New structures and new functions. *Trends Biochem. Sci.* **21,** 375–382.

Lupas, A., Vandyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252,** 1162–1164.

MacPhee, C. E., and Woolfson, D. N. (2004). Engineered and designed peptide-based fibrous biomaterials. *Curr. Opin. Solid State Mat. Sci.* **8,** 141–149.

McClain, D. L., Gurnon, D. G., and Oakley, M. G. (2002). Importance of potential interhelical salt-bridges involving interior residues for coiled-coil stability and quaternary structure. *J. Mol. Biol.* **324,** 257–270.

McClain, D. L., Woods, H. L., and Oakley, M. G. (2001). Design and characterization of a heterodimeric coiled coil that forms exclusively with an antiparallel relative helix orientation. *J. Am. Chem. Soc.* **123,** 3151–3152.

McLachlan, A. D., and Stewart, M. (1975). Tropomyosin coiled-coil interactions – Evidence for an unstaggered structure. *J. Mol. Biol.* **98,** 293–304.

McLachlan, A. D., Stewart, M., and Smillie, L. B. (1975). Sequence repeats in alpha-tropomyosin. *J. Mol. Biol.* **98,** 281–291.

Meier, M., Lustig, A., Aebi, U., and Burkhard, P. (2002). Removing an interhelical salt bridge abolishes coiled-coil formation in a *de novo* designed peptide. *J. Struct. Biol.* **137,** 65–72.

Melnik, T. N., Villard, V., Vasiliev, V., Corradin, G., Kajava, A. V., and Potekhin, S. A. (2003). Shift of fibril-forming ability of the designed alpha-helical coiled-coil peptides into the physiological pH region. *Prot. Eng.* **16,** 1125–1130.

Michnick, S. W. (2001). Exploring protein interactions by interaction-induced folding of proteins from complementary peptide fragments. *Curr. Opin. Struct. Biol.* **11,** 472–477.

Micklatcher, C., and Chmielewski, J. (1999). Helical peptide and protein design. *Curr. Opin. Chem. Biol.* **3,** 724–729.

Mittl, P. R. E., Deillon, C., Sargent, D., Liu, N. K., Klauser, S., Thomas, R. M., Gutte, B., and Grutter, M. G. (2000). The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. *Proc. Natl. Acad. Sci. USA* **97,** 2562–2566.

Moitra, J., Szilak, L., Krylov, D., and Vinson, C. (1997). Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **36,** 12567–12573.

Moll, J. R., Ruvinov, S. B., Pastan, I., and Vinson, C. (2001). Designed heterodimerizing leucine zippers with a range of pIs and stabilities up to 10(-15) M. *Prot. Sci.* **10,** 649–655.

Monera, O. D., Kay, C. M., and Hodges, R. S. (1994). Electrostatic interactions control the parallel and antiparallel orientation of alpha-helical chains in 2-stranded alpha-helical coiled-coils. *Biochemistry* **33,** 3862–3871.

Monera, O. D., Sonnichsen, F. D., Hicks, L., Kay, C. M., and Hodges, R. S. (1996a). The relative positions of alanine residues in the hydrophobic core control the formation of two-stranded or four-stranded alpha-helical coiled-coils. *Prot. Eng.* **9,** 353–363.

Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993). Comparison of antiparallel and parallel 2-stranded alpha-helical coiled-coils – design, synthesis, and characterization. *J. Biol. Chem.* **268,** 19218–19227.

Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M., and Hodges, R. S. (1996b). Formation of parallel and antiparallel coiled-coils controlled by the relative positions of alanine residues in the hydrophobic core. *J. Biol. Chem.* **271,** 3995–4001.

Myszka, D. G., and Chaiken, I. M. (1994). Design and characterization of an intramolecular antiparallel coiled-coil peptide. *Biochemistry* **33,** 2363–2372.

Nautiyal, S., and Alber, T. (1999). Crystal structure of a designed, thermostable, heterotrimeric coiled coil. *Prot. Sci.* **8,** 84–90.

Nautiyal, S., Woolfson, D. N., King, D. S., and Alber, T. (1995). A designed heterotrimeric coiled-coil. *Biochemistry* **34,** 11645–11651.

O'Neil, K. T., and DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids [published erratum appears in *Science* 1991, Aug 30; **253**(5023), 952]. *Science* **250,** 646–651.

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254,** 539–544.

O'Shea, E. K., Lumb, K. J., and Kim, P. S. (1993). Peptide velcro – design of a heterodimeric coiled-coil. *Curr. Biol.* **3,** 658–667.

O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* **68,** 699–708.

O'Shea, E. K., Rutkowski, R., Stafford, W. d., and Kim, P. S. (1989). Preferential heterodimer formation by isolated leucine zippers from fos and jun. *Science* **245,** 646–648.

Oakley, M. G., and Hollenbeck, J. J. (2001). The design of antiparallel coiled coils. *Curr. Opin. Struct. Biol.* **11,** 450–457.

Oakley, M. G., and Kim, P. S. (1998). A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry* **37,** 12603–12610.

Oas, T. G., and Endow, S. A. (1994). Springs and hinges – dynamic coiled coils and discontinuities. *Trends Biochem. Sci.* **19,** 51–54.

Pandya, M. J., Cerasoli, E., Joseph, A., Stoneman, R. G., Waite, E., Woolfson, D. N. (2004). Sequence and structural duality: Designing peptides to adopt two stable conformations. *J. Am. Chem. Soc.* **126,** 17016–17024.

Pandya, M. J., Spooner, G. M., Sunde, M., Thorpe, J. R., Rodger, A., and Woolfson, D. N. (2000). Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry* **39,** 8728–8734.

Parry, D. A. D. (1982). Coiled coils in α-helix containing proteins: Analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled coils in other proteins. *Biosci. Rep.* **2,** 1017–1024.

Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D., and Tirrell, D. A. (1998). Reversible hydrogels from self-assembling artificial proteins. *Science* **281,** 389–392.

Potekhin, S. A., Melnik, T. N., Popov, V., Lanina, N. F., Vazina, A. A., Rigler, P., Verdini, A. S., Corradin, G., and Kajava, A. V. (2001). *De novo* design of fibrils made of short alpha-helical coiled coil peptides. *Chem. Biol.* **8,** 1025–1032.

Rabindran, S. K., Haroun, R. I., Clos, J., Wisniewski, J., and Wu, C. (1993). Regulation of heat shock factor trimer formation: Role of a conserved leucine zipper. *Science* **259,** 230–234.

Ryadnov, M. G., Ceyhan, B., Niemeyer, C. M., and Woolfson, D. N. (2003). ''Belt and braces'': A peptide-based linker system of *de novo* design. *J. Am. Chem. Soc.* **125,** 9388–9394.

Ryadnov, M. G., and Woolfson, D. N. (2003a). Engineering the morphology of a self-assembling protein fibre. *Nature Mat.* **2,** 329–332.

Ryadnov, M. G., and Woolfson, D. N. (2003b). Introducing branches into a self-assembling peptide fiber. *Angew. Chem. Int. Ed.* **42,** 3021–3023.

Ryadnov, M. G., and Woolfson, D. N. (2004). Fiber recruiting peptides: Noncovalent decoration of an engineered protein scaffold. *J. Am. Chem. Soc.* **126,** 7454–7455.

Schnarr, N. A., and Kennan, A. J. (2001). Coiled-coil formation governed by unnatural hydrophobic core side chains. *J. Am. Chem. Soc.* **123,** 11081–11082.

Schnarr, N. A., and Kennan, A. J. (2002). Peptide tic-tac-toe: Heterotrimeric coiled-coil specificity from steric matching of multiple hydrophobic side chains. *J. Am. Chem. Soc.* **124,** 9779–9783.

Schnarr, N. A., and Kennan, A. J. (2003). pH-triggered strand exchange in coiled-coil heterotrimers. *J. Am. Chem. Soc.* **125,** 6364–6365.

Schneider, J. P., Lombardi, A., and DeGrado, W. F. (1998). Analysis and design of three-stranded coiled coils and three- helix bundles. *Fold. Des.* **3,** R29–R40.

Sharma, V. A., Logan, J., King, D. S., White, R., and Alber, T. (1998). Sequence-based design of a peptide probe for the APC tumor suppressor protein. *Curr. Biol.* **8,** 823–830.

Sia, S. K., and Kim, P. S. (2001). A designed protein with packing between left-handed and right-handed helices. *Biochemistry* **40,** 8981–8989.

Steinmetz, M. O., Stock, A., Schulthess, T., Landwehr, R., Lustig, A., Faix, J., Gerisch, G., Aebi, U., and Kammerer, R. A. (1998). A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *EMBO J.* **17,** 1883–1891.

Stevens, M. M., Flynn, N. T., Wang, C., Tirrell, D. A., and Langer, R. (2004). Coiled-coil peptide-based assembly of gold nanoparticles. *Adv. Mat.* **16,** 915–918.

Su, J. Y., Hodges, R. S., and Kay, C. M. (1994). Effect of chain-length on the formation and stability of synthetic alpha-helical coiled coils. *Biochemistry* **33,** 15501–15510.

Suzuki, N., and Fujii, I. (1999). Optimization of the loop length for folding of a helix-loop-helix peptide. *Tetrahedron Letts.* **40,** 6013–6017.

Talbot, J. A., and Hodges, R. S. (1982). Tropomyosin – A model protein for studying coiled-coil and alpha-helix stabilization. *Acc. Chem. Res.* **15,** 224–230.

Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000). Effects of side-chain characteristics on stability and oligomerization state of a *de novo* – designed model coiled-coil: 20 amino acid substitutions in position ''d.'' *J. Mol. Biol.* **300,** 377–402.

Vinson, C. R., Hai, J. W., and Boyd, S. M. (1993). Dimerization specificity of the leucine zipper-containing BZIP motif in DNA binding — Prediction and rational design. *Genes Dev.* **7,** 1047–1058.

Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999). The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Prot. Sci.* **8,** 2312–2329.

Walshaw, J., and Woolfson, D. N. (2001a). Open-and-shut cases in coiled-coil assembly: Alpha-sheets and alpha-cylinders. *Prot. Sci.* **10,** 668–673.

Walshaw, J., and Woolfson, D. N. (2001b). SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307,** 1427–1450.

Walshaw, J., and Woolfson, D. N. (2003). Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J. Struct. Biol.* **144,** 349–361.

Wang, C., Stewart, R. J., and Kopecek, J. (1999). Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* **397,** 417–420.

Wolf, Y. I., Grishin, N. V., and Koonin, E. V. (2000). Estimating the number of protein folds and families from complete genome data. *J. Mol. Biol.* **299,** 897–905.

Woolfson, D. N. (2001). Core-directed protein design. *Curr. Opin. Struct. Biol.* **11,** 464–471.

Woolfson, D. N., and Alber, T. (1995). Predicting oligomerization states of coiled coils. *Prot. Sci.* **4,** 1596–1607.

Yeates, T. O., and Padilla, J. E. (2002). Designing supramolecular protein assemblies. *Curr. Opin. Struct. Biol.* **12,** 464–470.

Yoder, N. C., and Kumar, K. (2002). Fluorinated amino acids in protein design and engineering. *Chem. Soc. Rev.* **31,** 335–341.

Zhou, M., Bentley, D., and Ghosh, I. (2004). Helical supramolecules and fibers utilizing leucine zipper-displaying dendrimers. *J. Am. Chem. Soc.* **126,** 734–735.

Zimenkov, Y., Conticello, V. P., Guo, L., and Thiyagarajan, P. (2004). Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif. *Tetrahedron.* **60,** 7237–7246.

MICRODISSECTION OF THE SEQUENCE AND STRUCTURE OF INTERMEDIATE FILAMENT CHAINS

By DAVID A. D. PARRY

Institute of Fundamental Sciences, Massey University,
Palmerston North 5301, New Zealand

I. Introduction 114
II. Sequence Regularities and Characteristics, and Their Effects on Secondary and Tertiary Structure 116
A. Head Domain 116
B. Segment 1A 120
C. Linker L1 123
D. Segment 1B 125
E. Linker L12 127
F. Segment 2A 128
G. Linker L2 128
H. Segment 2B 129
I. Tail Domain 132
J. Features in Head and Tail Domains 134
III. Summary 136
References 137

Abstract

A large number of intermediate filament (IF) chains have now been sequenced. From these data, it has been possible to deduce the main elements of the secondary structure, especially those lying within the central rod domain of the molecule. These conclusions, allied to results obtained from crosslinking studies, have shown that at least four unique but related structures are adopted by the class of structures known generically as intermediate filaments: (1) epidermal and reduced trichocyte keratin; (2) oxidized trichocyte keratin; (3) desmin, vimentin, neurofilaments, and related Type III and IV proteins; and (4) lamin molecules. It would be expected that local differences in sequences of the proteins in these four groups would occur, and that this would ultimately relate to assembly. Site-directed mutagenesis and theoretical methods have now made it possible to investigate these ideas further. In particular, new data have been obtained that allow the role played by some individual amino acids or a short stretch of sequence to be determined. Among the observations catalogued here are the key residues involved in intra- and interchain ionic interactions, as well as those involved in stabilizing some modes of

DOI: 10.1016/S0065-3233(04)70005-4
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

molecular aggregation; the structure and role of subdomains in the head and tail domains; the repeat sequences occurring along the length of the chain and their structural significance; trigger motifs in coiled-coil segments; and helix initiation and termination motifs that terminate the rod domain. Much more remains to be done, not least of which is gaining an increased understanding of the many subtle differences that exist between different IF chains at the sequence level.

I. Introduction

Over the past 25 years, a wealth of amino acid sequence data have been derived for the class of proteins known generically as *intermediate filaments* (IFs). These include: the Type I and Type II chains characteristic of trichocyte and epidermal keratins; the Type III chains of desmin, vimentin, glial fibrillary acidic protein, and peripherin; the Type IV chains in neurofilaments and of α-internexin; the Type V chains of nuclear lamins; and, finally, the Type VI chains of nestin (and possibly synemin and paranemin). There is some debate on the classification of nestin as a Type VI IF protein, and Shaw (1998) has suggested that it should be designated within the Type IV group on the basis of gene structure. Other IF proteins include phakinin and filensin of beaded filaments in lens fiber cells. While all of these proteins are intracellular ones (though occurring at both cytoplasmic and nuclear sites), there is at least one other group that occurs extracellularly: these proteins are produced in the gland thread cells of the hagfish (Downing *et al.*, 1981; Koch *et al.*, 1991, 1994, 1995). Furthermore, intermediate filaments also occur widely in invertebrates (see, for example, Riemer *et al.*, 1998).

Some IF protein chains are able to homopolymerize *in vivo* to form structurally viable and functional IFs. These include desmin, vimentin, peripherin, glial fibrillary acidic protein, NF-L, nestin, and phakinin. Some other chains only form IFs if they have an appropriate copolymerization partner. These include the keratins (which form Type I/Type II molecules: Hatzfeld and Weber, 1990; Herrling and Sparrow, 1991; Parry *et al.*, 1985; Steinert, 1990), nestin (which copolymerizes with vimentin and α-internexin: Steinert *et al.*, 1999a,b), paranemin and synemin (with desmin or vimentin: Bilak *et al.*, 1998), filensin (with phakinin: Goulielmos *et al.*, 1996), and both the medium (NF-M) and heavy (NF- H) neurofilament chains (with NF-L: Ching and Liem, 1993; Lee *et al.*, 1993). Some IF proteins are able to form both homo- and heteropolymers depending on circumstances: desmin, vimentin, peripherin, glial fibrillary acidic protein, α-internexin, nestin, phakinin, and NF-L (Herrmann and Aebi, 2000;

Parry and Steinert, 1999; Steinert *et al.*, 1999a). There is increasing evidence that the copolymerized molecules may be located preferentially on the surface of IFs and that the core comprises predominantly homodimers of a more abundant chain type (Goulielmos *et al.*, 1996; Herrmann and Aebi, 2000).

The primary structure data have provided a strong basis for both theoreticians and experimentalists to gain an increased understanding of the conformation of an individual IF chain, its resulting molecular structure, and its mode of assembly to form viable IFs. All of these IF chains are homologous to one another to varying degrees over a central region, and distinct subdomains have been described. In particular, there are four regions with a heptad substructure that are predicted to form coiled coils; these are designated 1A, 1B, 2A, and 2B. They are separated by short noncoiled coil regions lacking a heptad repeat, known as linkers (L1, L12, and L2 respectively; Fig. 1a). In contrast, the end domains of

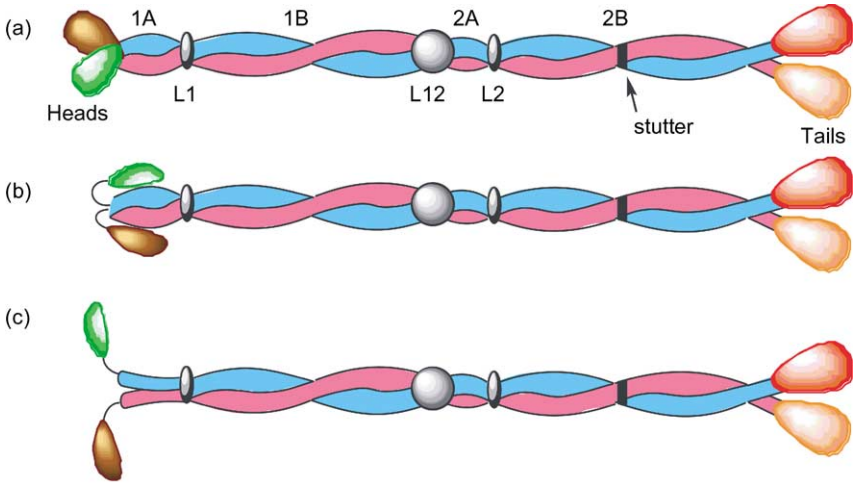


FIG. 1. Schematic diagram of (a) an intermediate filament heterodimer with coiled-coil domains 1A, 1B, 2A, and 2B, and noncoiled-coil connecting linkers L1, L12, and L2. A stutter occurs in the heptad substructure at a point close to the center of segment 2B. The N-terminal ''globular'' domains (green for Type I and brown for Type II chains) are termed the heads, and the C-terminal domains (red for Type I and orange for Type II chains) are designated the tails. In (b), the heads are shown folded back over the rod domain, where it is believed that this will stabilize segment 1A. In (c), the heads are shown away from the body of the rod domain and in a position where they can interact more easily with other cellular entities. As a consequence, segment 1A may become destabilized and hence unwind to form two separate α-helical strands.

different chain types display considerable variability in both size and chemical character. In turn, the molecule contains two parallel chains lying in axial register; these form a central rodlike domain with a coiled-coil structure. The rod separates the ''globular'' head and tail domains, and in so doing forms a tripartite molecule. In simple terms, the rod domains dominate (though not absolutely) in the assembly of molecules to form the core of the filaments. A significant fraction of the end domains is thus disposed in outer positions on the filament surface, where they can interact optimally either with one another or with other cellular entities for functional and/or structural reasons. A number of reviews are available regarding IF structure, assembly, and function (see, for example, Herrmann *et al.*, 2003; Parry, 1997; Parry and Steinert 1995, 1999; Hermann and Aebi, 1999a). These will provide the reader with additional and useful background material to the subject matter that is the theme of this Chapter.

The purpose of this review is to summarize details of both the specific structures and roles which individual amino acids, short stretches of amino acids, and subdomains play *in vivo*. In some cases, the pertinent sequence identified as significant on the basis (perhaps) of high sequence conservation or regularity remains largely uninterpretable in detail, though often strong indications may exist as to its likely structure/function. In other instances, there is a high degree of very specific information at hand; this will be chronicled here individually for each residue, sequence, or segment of structure, as the case may be. It is believed that this compendium is the first of its type and that it will, hopefully, provide a useful insight to those in the field as to the gaps that exist in our present knowledge.

II. Sequence Regularities and Characteristics, and Their Effects on Secondary and Tertiary Structure

A. *Head Domain*

To facilitate easy access to information, this section has been subdivided (in most cases, purely on the basis of chain type). In a number of instances, however, published data refer generically to both head and tail domains; in order not to repeat the information in Sections II.A and II.I, a separate section (Section II.J) has been written to cover these features. They include: (1) the covalent binding sites in head and tail domains that are involved in crosslinks with other proteins in the cell; and (2) post-translational modifications and their structural/functional effects *in vivo*.

1. Epidermal Keratins (Type I/Type II)

The head domain in all epidermal keratins displays a highly developed substructure (Fig. 2a: Steinert *et al.*, 1985). The region immediately N-terminal to the beginning of the rod domain is designated the Homologous 1 (H1) subdomain and is characterized by a well-conserved sequence of seven residues in Type I chains and 36 residues in Type II chainsEquivalent regions are found in the trichocyte keratins (Section II.A.2). The consensus sequences are given in Parry and Steinert (1999). While there is little knowledge of the precise structure of the H1 subdomain, it has been shown experimentally that in Type II chains its functional role is to specify molecular aggregation at the two- to four-molecule level (Hatzfeld and Burba, 1994; Steinert and Parry, 1993; Steinert *et al.*, 1993a). N-terminal to H1 is subdomain Variable 1 (V1), which has a variable length and sequence but is nonetheless characterized by a high content of glycine and serine residues. V1 has a length in the range 65 to 140 residues in Type I chains and 20 to 160 residues in Type II chains. Finally, the most N-terminal region is called the End (E) subdomain, and this is often relatively basic in character. It is short in Type I chains (0–10 residues) but generally longer in Type II chains (5–70 residues). Little is known about the overall conformation of head domains in epidermal keratins, though a good working model displaying a glycine loop structure for subdomain V1 is available (Korge *et al.*, 1992a, 1992b; Steinert *et al.*, 1991). In essence, runs of glycine-rich sequence are anchored by interactions between aromatic and/or large apolar residues (Fig. 2b). On theoretical (Conway *et al.*, 1989) and experimental grounds (Mack *et al.*, 1988), such a conformation has been shown to be very flexible. The glycine loop model also allows deletions and insertions to occur, as has been characterized in detail for the V2 subdomain in human K10 (see Section II.I). The structure also permits variations in glycine loop size to occur without impinging adversely on the remaining structure.

2. Trichocyte Keratins (Type I/Type II)

The head domain in trichocyte (or hair) keratins has been characterized by Parry and North (1998) and Parry *et al.* (2002) and shown to consist of two domains—a basic one (NB) at the N-terminal end of the head domain (27 and 70 residues long, respectively, for Type I and Type II chains) and an acidic one (NA) lying between the rod and NB. The latter is 29 and 34 residues long for Type I and Type II chains, respectively. Importantly, however, NA is homologous to H1 in the respective chain types of epidermal keratins. There can therefore be little, if any, doubt that the role of these regions in trichocyte and epidermal keratins is

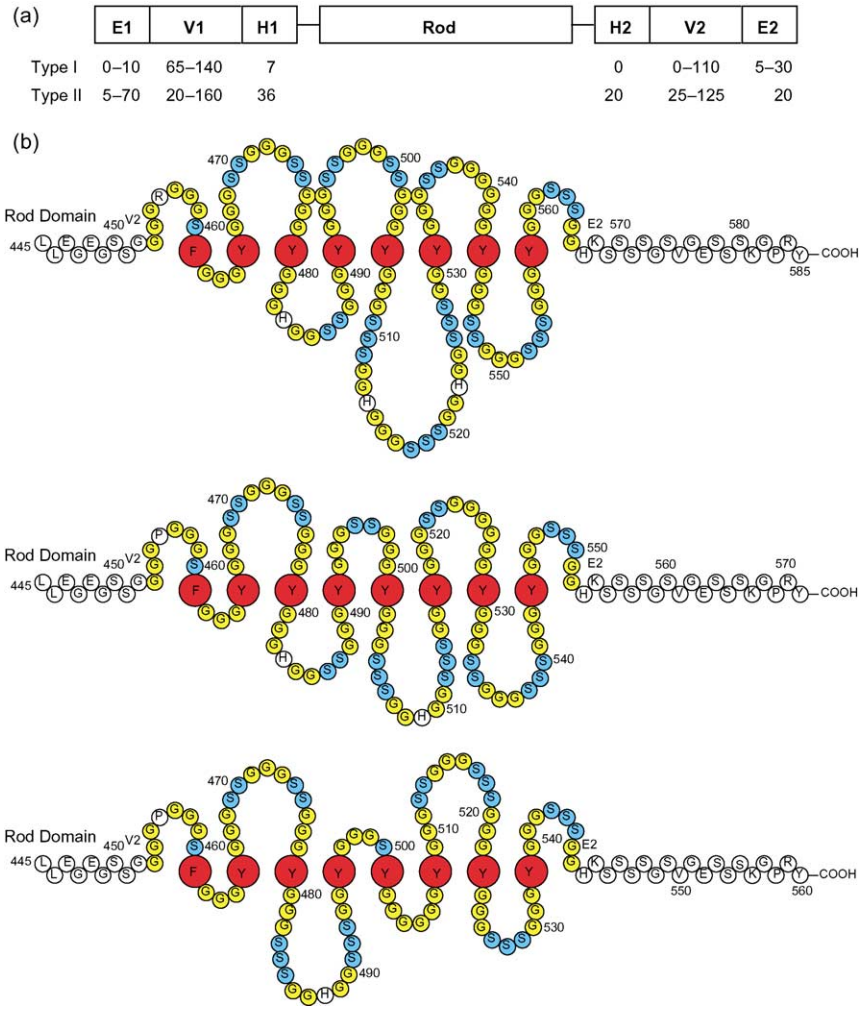


FIG. 2. (a) Schematic diagram of the subdomain structure in epidermal keratin chains. Homologous (*H*), variable glycine-serine-rich (*V*) and generally basic (*E*) end domains are present and these have bilateral symmetry about the central rod domain. The H subdomains have a special role in stabilizing oligomers (about 2–4 molecules) during the early stages of assembly to form filaments. The numbers of residues in each subdomain are indicated below for both the Type I and Type II chains. (b) The sequences of the V2 subdomain in the human K10 chain show marked differences in allelic variants that have been characterized. Deletions and insertions are seen in the loop structure anchored and stabilized by interactions between the aromatic residues tyrosine and phenylalanine. Glycine and serine residues dominate the loops, which are believed to be highly flexible structures. Redrawn from Parry and Steinert (1995) and based on an original by Korge *et al.* (1992a).

identical (i.e., they stabilize assembly at the two- to four-molecule level). NA and NB both have clear-cut amino acid composition characteristics, which differ considerably between chain types.

The Type I head domain structure is not easily defined. Starting from the N-terminal end of the chain, residues 1 through 11 lack charged amino acids. Residues 12 through 18 are partially conserved and contain both serine and arginine in greater amounts than expected on average. Residues 19 through 47 lack charged amino acids but cannot be usefully subdivided further. After a single glutamic acid (residue 48), there is a region (residues 49–55) directly homologous to H1 in the Type I epidermal chain (Fig. 3a). Overall, residues 1–11 and 19–47 have an underlying nonapeptide repeat of the form C/F–F/N–X–F/P–C–L–P–G–S. There are no indications at present as to the conformation adopted by this quasi-motif.

In the Type II head domains, the sequences display little homology over the first 12 residues but there is considerable consensus over the next 24 (residues 13–36). This sequence is R–A–F–S–C–V–S–A–C–G–P–R–P–

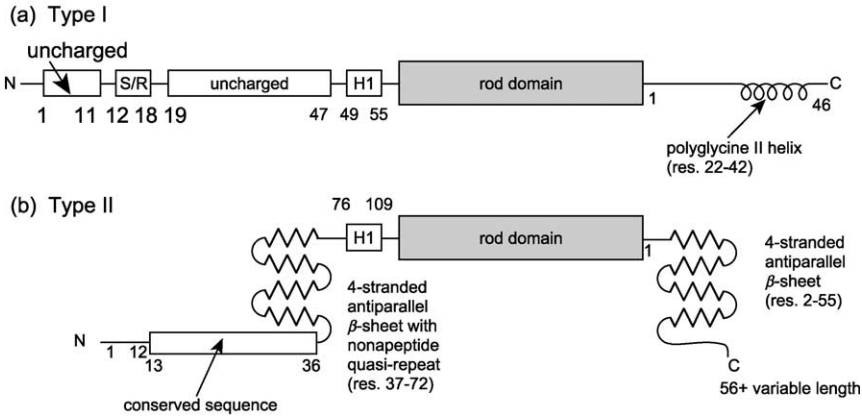


FIG. 3. Schematic diagram of (a) Type I and (b) Type II trichocyte keratin chains illustrating zones of particular character in the sequences and the proposed secondary structure in both the head and tail domains. The head domain in both the Type I and Type II chains contains a conserved H1 subdomain believed to be important in stabilizing the structure at the 2- to 4-molecule level of assembly. The Type I chain also contains two segments lacking charged residues that enclose a short region rich in serine and arginine residues. The Type II chain contains a four-stranded antiparallel β-sheet with an underlying but poorly conserved nonapeptide repeat, as well as 24-residue stretch of sequence displaying high conservation. Its role is unknown. The tail domains are predicted to contain a poly-glycine II helix and a four-stranded antiparallel β-sheet respectively in the Type I and Type II chains.

G–R–C–C–I–T/S–A–A–P–Y–R. C-terminal to this region, there are four contiguous but imperfect nonapeptide repeats (residues 37–72) of the form $(G–G–F–G–Y–R–S–X–G)_4$. Note that this repeat is different from the one described in the Type I head domain. This region is predicted to constitute a four-stranded antiparallel β-sheet that folds back over segment 1A and interacts with it under some circumstances, or which interacts with an equivalent β-sheet in another molecule in a different IF in other circumstances (Parry *et al.*, 2002). Evidence in support of part of this premise is the fact that disulphide crosslinks have been induced in mouse hair between cys 99 in the Type II head domain and cys 6 in segment 1A of the Type II chain, and between cys 75 in the Type II head domain and cys 30 in segment 1A of the Type I chain (Parry *et al.*, 2002). Following a V–G–G tripeptide, the NA domain (residues 76–109) is completed with a consensus sequence of P–S–P–P–C–I–T–T–V–S–V–N–E–S–L–L–T–P–L–N–L–E–I–D–P–N–A–Q–C–V–K–Q/H–E–E. As noted before, this is homologous to H1 in the epidermal type II chain (Fig. 3b).

3. Desmin, Vimentin, Glial Fibrillary Acidic Protein, and Peripherin (Type III) and Neurofilaments and α-Internexin (Type IV)

From a wide range of species, the head domain in vimentin, desmin, peripherin, and Type IV IF protein contains a conserved nonapeptide of the form S–S–Y–R–R–T–F–G–G. This is required for regular assembly to proceed (Herrmann *et al.*, 1992). The nonapeptide sequence is missing, however, in glial fibrillary protein but two other motifs have been identified (Ralton *et al.*, 1994). Herrmann and Aebi (1998b) have pointed out that, in addition to the conserved nonapeptide, there are six aromatic residues roughly equally spaced in the head domain, as well as 12 arginine residues. Overall, the head domain has a strongly hydrophilic character. A possible conformation would be one akin to that proposed for the glycine-serine-rich V domains in epidermal keratin. In that structure, and possibly here also for Type III head domains, a ''glycine loop'' type of structure may exist with loops being anchored through strong aromatic interactions (Fig. 2b).

B. Segment 1A

Segment 1A is defined in large part by its underlying heptad repeat, which has the form $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, where positions *a* and *d* are largely occupied by apolar residues. Such a motif is characteristic of an α-helical conformation that aggregates with others to form a multistranded coiled-coil rope. In the case of IF molecules, there are two strands only, which are aligned parallel to one another and in axial register (see, for example,

Parry *et al.*, 1985 and the crosslinking data of Steinert *et al.*, 1993a,b,c, 1999a).

The length of segment 1A is invariant (35 residues) in all IF chains. It is characterized by a number of unusual features and is clearly differentiated from other coiled-coil segments (1B, 2A, and 2B) in IF proteins. First, segment 1A contains more highly conserved residues across all IF chain types than any other coiled-coil segment in the rod domain: L7, N8, A12, V18, L21, E22, and N25 (Smith *et al.*, 2002). This points to particularly important roles for segment 1A, some of which are described below. (One particular role—that of a head-to-tail overlap between similarly directed molecules—is discussed in Section II.H). Second, towards the N-terminal end of 1A, there is a very highly conserved stretch of sequence known as the *helix initiation motif* that is rich in apolar and aromatic residues. It encompasses residues 1 through 22 and has a consensus sequence of K/R–X–T/Q–M/I/L–K/Q–X–L–N–D–R–F–A–S–F/Y–I–D/E–K–V–R–F–L–E (Parry and Steinert, 1999; Smith *et al.*, 2002). Third, at the C-terminal end of this region, there is a potential interchain ionic interaction that is conserved across all chain types (Smith *et al.*, 2002). This occurs between a basic residue in the *g* position (residue 17) and an acidic residue in the *e* position (residue 22). Fourth, two of the three regions in the rod domain displaying the greatest hydropathy are located in segment 1A (residues 12–21 and 27–33). The third region lies at the end of segment 2B (Section II.I). Fifth, there is a single potential intrachain ionic interaction that is conserved across all chain types. It is of the type i to $i + 3$ and occurs between an acidic residue in position 16 and an arginine residue in position 19 (Smith *et al.*, 2002).

Next, although the heptad repeat is a well-conserved feature of segment 1A, there are unique features present that are absent in other coiled-coils in α-fibrous proteins, including the others in IF proteins. This is manifested in an unusually high apolar residue content in positions *a* and *d*, especially the latter (76% in *a* and 92% in *d* compared to average values of 73% and 76%, respectively). Also, the leucine residue content in position *d* (68%) is very much greater than normal (44%). The most significant difference, however, lies in the distribution of the amino acids in each of the heptad positions as compared to both an average two-stranded coiled coil and to segments 1B and 2B in particular (Smith *et al.*, 2002). These were listed by the authors as follows: in position *b*, K increases from 2.5% to 14.6%, R increases from 6.8% to 23.0%, and E decreases from 19.6% to 3.6%; in position *c*, R decreases from 10.4% to 2.9%; in position *e*, E increases from 22.1% to 31.8 %, and R decreases from 16.7% to 3.8%; in position *f*, E increases from 9.2% to 20.7%, D increases from 7.1% to 26.8%, and R decreases from 11.5% to 2.0%; in position *g*, E decreases

from 27.7% to 16.4%, R increases from 6.1% to 21.8%, and K increases from 8.4% to 28.0%. Thus, positions *c*, *e*, and *f* become more acidic/less basic and positions *b* and *g* become more basic/less acidic compared to two-stranded coiled-coils in general.

In the seventh difference, segment 1A lacks any significant periodicity in the linear distribution of its acidic or its basic residues. This is in direct contrast to these same residue groupings in both segment 1B and segment 2 (2A–L2–2B). In these latter cases, stabilizing ionic interactions arising from the regular disposition of oppositely charged residues are believed to be important in the molecular aggregation process. Segment 1A, in contrast, lacks such interactions. Finally, segment 1A represents one of the two mutation hotspots in IF proteins (Parry and Steinert, 1995, 1999). This again suggests an important structural/functional role for this region.

In order to study the residues likely to be important in stabilizing various modes of molecular aggregation, point mutations of charged residues were engineered along the length of the K5/K14 epidermal keratin chains. Aggregation characteristics of equimolar mixtures of wild-type and mutant chains (Mehrani *et al.*, 2001) were then studied. The data show that the conserved residue R10 in segment 1A was essential for stability and that D/N9 was also important in this regard. These two residues lay in close axial proximity in the A_{11} mode to conserved residues 4E and 6E in linker L2 that were shown to be essential for stability, and to residue D7 in linker L2 that also has an important role in providing stability. Likewise, the results have shown that the conserved residue D1 in segment 2A was important for stability, as were D/N3 (also in segment 2A) and conserved residue K31 in segment 1A. These two groups of residues (1A-9/10 with L2-4/6/7 and also 1A-31 with 2A-1/3) lay in close axial proximity in the A_{11} mode, and it was suggested that they are involved in making stabilizing intermolecular ionic interactions and hydrogen bonds (Mehrani *et al.*, 2001).

Twenty-two naturally occurring mutations observed in segment 1A of patients suffering from epidermolysis bullosa simplex (EBS) have been studied using a molecular dynamics approach (Smith *et al.*, 2004). This was a comprehensive and computationally exhaustive study that yielded a number of interesting results. For example, the conformational changes observed fell into five categories. In regards to the backbone structure, the mutation caused: (1) a local structural change only in the chain in which the mutation was located; (2) a structural change only in the partner chain, usually to the extent of one or two turns of α-helix; (3) local structural changes in both chains, but again only locally; (4) widespread conformational change; and (5) little obvious conformational

modification. Furthermore, the effect of mutations in the inner *a/d*, adjacent *e/g*, and outer positions *b/c/f* of the coiled coil were (surprisingly) directly comparable. The same mutation in different chains of the heterodimer did not necessarily give rise to the same structural change. Different mutations at the same site (e.g., 1A-10) can give rise to quite different structural changes, indicating that it is currently impossible to predict from first principles the extent of a structural change introduced by a specific mutation.

Unexpectedly, crystallographic studies of segment 1A revealed that the structure was not that of a two-stranded coiled coil but was instead characterized by separate α-helical strands (Strelkov *et al.*, 2002). Each individual α-helix had a coiled axis (radius of curvature, 8.4 nm) and it was easily shown that two such strands could be assembled computationally to form a coiled-coil molecule (radius, 0.5 nm; pitch length, 16.5 nm) that was virtually identical to that of GCN4, an archetypical coiled-coil structure (Smith *et al.*, 2002). The idea that segment 1A could form a coiled coil under some circumstances, but two individual α-helices under other conditions, was an exciting one that intimated a more dynamic role for the N-terminal end of the rod domain than had hitherto been assumed (Herrmann *et al.*, 2003; Parry *et al.*, 2002). The observation by Strelkov *et al.* (2001) that dimers formed only in the presence of the head domain can be interpreted in terms of segment 1A being able to unwind. Furthermore, it seems probable that the head domain stabilizes the dimeric form of segment 1A by folding back over it (Fig. 1b and c). Smith *et al.* (2002) calculated the stability of the α-helices in all coiled-coil segments in the rod domain of IF proteins, and were able to demonstrate that the α-helices in segment 1A were the most stable. This is consistent with them being able to exist as individual α-helices when segment 1A unwinds into its component strands.

C. Linker L1

Linker L1 is a short stretch of sequence that connects coiled-coil segments 1A and 1B. In contrast to these segments, L1 lacks both their heptad repeat and their α-helical structure, the only exception being that of the Type V lamins. The length of L1 is typically 9–16 residues for Type I chains, 10–14 residues for Type II, 8–11 residues for Type II, 9–10 residues for Type IV, and 9 residues for Type VI. The special case of lamin Type V chains is discussed below. In general, the L1 sequences from different chain types show very little similarity to one another over their central regions. However, limited homology does exist at both ends of L1. All L1

linkers have a high charged/apolar residue ratio (typically about 1.6), characteristic of an elongate conformation. Indeed, Parry *et al.* (2001) have shown that linker L1 is the most elongate rod domain element in terms of the average unit rise per residue. Crosslinking data show that the axially projected length of L1 is 14.84 h_{cc} (2.20 nm) in Type I and Type II keratins, 11.47 h_{cc} (1.70 nm) in Type III vimentin (Parry *et al.*, 2001), and 14.50 h_{cc} (2.15 nm) in Type IV α-internexin (Parry and Steinert, 1999), where h_{cc} is the unit rise per residue in a coiled-coil conformation (0.1485 nm). Within the experimental limitations, these values probably do not differ significantly from one another.

Unlike other IF chains, linker L1 in the lamins is predicted to be α-helical. Furthermore, the heptad repeat in segment 1A can be extended C-terminally, and then in segment 1B N-terminally, to encompass the entire region normally defined as L1. The heptad repeat does suffer a phase change at the region normally termed L1, and this corresponds to an insertion of one residue or a deletion of six residues in an otherwise continuous heptad repeat. This type of heptad interruption is called a *skip* and is structurally equivalent to a pair of closely spaced stutters (three residue deletions). This results in a local unwinding of the coiled coil so that the strands probably coil around one another locally in a right-handed manner (Brown *et al.*, 1996; Burkhard *et al.*, 2001; Kühnel *et al.*, 2004; Özbek *et al.*, 2004; Strelkov and Burkhard, 2002; Stetefeld *et al.*, 2000). Thus, in lamins, segment 1 (1A + L1 + 1B) effectively adopts a continuous coiled-coil structure over its entire length, and in this respect Type V lamins differ significantly from all other IF chain types.

Excluding the lamins, linker L1 is the most flexible rod domain region. This was initially shown by Conway *et al.* (1989) and more recently by Smith *et al.* (2002), who confirmed that L1 did indeed have the highest flexibility index of any rod domain element in chain types I–IV. It would appear that flexibility for linker L1 is an important structural feature of IF molecules. Further proof of this point comes from the elegant work of Herrmann and Aebi (1998b) and Herrmann *et al.* (1999). A variety of mutants of L1 were prepared, including one in which two additional α-helix–favoring alanine residues were inserted and one residue was mutated to an alanine. This maintained the heptad phasing from segment 1A through to segment 1B and gave the sequence a greater probability of forming an α-helical conformation. Assembly of this mutant, however, was adversely affected and only short and rather irregular aggregates were formed. Also, a mutant of L1 in which one residue was changed to a non-α-helix–favoring proline residue and two additional prolines were added gave rise to normal filament formation. It therefore is important for L1 not to form a regular α-helical coiled-coil

structure, and some flexibility is clearly crucial for correct assembly and filament formation.

A mutation in L1 of particular interest has been reported by Mücke *et al.* (2004). In this case, the first residue in the linker in human vimentin has been altered from a lysine to a cysteine residue. At 37 °C, the K1C mutated protein assembles normally into unit-length filaments (ULFs) within about 10 seconds and into normal IFs after about an hour. When assembly is carried out at room temperature, ULFs are formed in the normal manner. However, elongation and assembly of the ULFs to form intact IFs was impeded significantly. Thus, the K1C mutation has a significant heat-sensitive effect, not on the formation of the important ULF intermediary, but on its ability to aggregate axially with others to form viable IFs. The rationale for this observation is not yet known, but several explanations exist. The lysine residue in the wild-type protein might have been involved in a salt bridge and the cysteine might be able to form a disulphide bond. It also needs to be remembered that apolar interactions are stronger at high temperatures.

In many respects, the linker L1 can be likened to a flexible hinge, a role consistent with a swinging head model proposed for IFs. In this model, segment 1A can (under appropriate conditions) split into two separate α-helical strands, thereby maximizing the range of movement of the head domains and hence increasing their capacity to interact with various other cellular entities (Parry *et al.*, 2002).

D. Segment 1B

Segment 1B has an uninterrupted heptad substructure (101 residues long) in all IFs except for Type V lamin and some invertebrate IF chains (Fig. 4). In these latter cases, segment 1B contains a 42-residue (six heptad) insertion located about 40 residues from its N-terminus (Fisher *et al.*, 1986; McKeon *et al.*, 1986; Parry *et al.*, 1986). When a comparison is made of all IF sequences, it is apparent that three residues are virtually conserved across the entire spectrum indicating a special role, probably structural rather than functional. These residues are L74, L81, and E95 (Smith *et al.*, 2002). In addition, there is a strong possibility of a conserved interchain ionic interaction in all IF molecules between a conserved acidic residue in a *g* position of one chain (D/E84) and a conserved basic residue in an *e* position of a second chain (K/R89) (Smith *et al.*, 2002). In precisely this same region (and almost certainly not by chance), a potential trigger motif has been recognized for epidermal keratin K5/K14 defined by residues 79–91 (Wu *et al.*, 2000). This has a

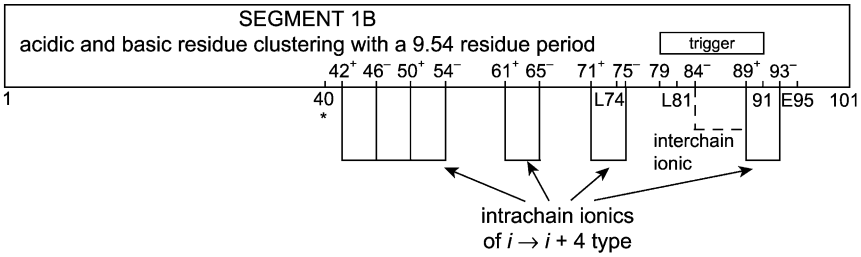


FIG. 4. Schematic diagram of segment 1B in intermediate filament chains showing generally conserved features. These include the highly conserved residues L74, L81, and E95; the conserved intra- and interchain ionic interactions represented by solid and dotted lines, respectively; the trigger motif (residues 79–91) that acts as a particularly stable region that ''nucleates'' coiled-coil formation; the site at residue 40 (indicated by an asterisk) of a six heptad insertion in lamin molecules. The entire segment displays a regular disposition of acidic and basic residues, each with a period of about 9.54 residues. These periods are approximately out of phase with one another.

sequence of D–A–L–M–D–E–I–N–F–M–K–M–F (K5 chain) and E–S–L–K–E–E–L–A–Y–L–K–K–N (K14 chain). Trigger motifs are believed to be especially important in providing stability for a coiled coil and, as a consequence, may act as a nucleating point for subsequent coiled-coil formation. In the case of cortexillin (Burkhard *et al.*, 2000), stability was achieved through a network of intra- and interchain ionic interactions.

Intrahelical ionic interactions are also important in stabilizing the individual α-helices in the IF molecule. There are a large number of conserved but oppositely charged residue pairs in segment 1B. These are all of the type i to $i + 4$, and are as follows: 42K/R-46E, 46E-50R, 50R-54E, 61R/K-65D, 71R/K-75E, and 89K/R-93E/D (Fig. 4; Smith *et al.*, 2002). In all but one case, the ionic bonds form between a basic residue in position i and an acidic residue in position $i + 4$.

The acidic and the basic residues are not randomly distributed in segment 1B. An analysis of their axial distributions using Fourier transform methods has revealed that both groupings display a very significant period of about 9.54 residues (1.42 nm). The periods are roughly, but not exactly, out of phase with one another (Crewther *et al.*, 1983; McLachlan and Stewart, 1982; Parry *et al.*, 1977). The structural rationale for the alternating bands of positive and negative charge is self-evident: 1B segments in different molecules will be in a position to interact strongly with one another if their zones of opposite charge become axially aligned (see Chapter 2, Fig. 6). Clearly, a family of closely related alignments is possible (Fraser *et al.*, 1985, 1986). Detailed molecular modeling of the interactions between 1B segments in vimentin (Smith *et al.*, 2003)

confirms that factors governing molecular alignment in the A_{11} mode lie not only in charged interactions between 1B segments, but also with interactions between various other parts of the molecule (note, for example, the role of subdomain H1 in epidermal keratins; see Section II.A.1).

X-ray studies on oriented films formed from isolated 1B segments from trichocyte keratin (wool) have revealed a typical coiled-coil diffraction pattern and axial period of 16 nm (Suzuki *et al.*, 1973). The length of segment 1B is about 15 nm (101 × 0.1485 nm); hence, it was proposed that the relative stagger between oppositely directed 1B segments must be small (<15 residues). From crosslinking data subsequently obtained, the relative stagger was shown to lie in the 15 to 20 residue range (Wang *et al.*, 2000), a value in close agreement with the estimate derived from *in vitro* X-ray diffraction experiments.

E. Linker L12

Linker L12, a region lacking a heptad repeat, lies at the center of the rod domain. Its length varies, though not by a great deal, from one chain type to another. It is 16, 17, 16–18, 15–22, 19, and 19–21 residues long, respectively, for the vast majority of the Type I, II, III, IV, V, and VI chains characterized thus far. In terms of physical projected length in the filament as determined from crosslinking studies (Parry *et al.*, 2001), this corresponds to $12.57h_{cc}$ (1.87 nm) for trichocyte and epidermal keratins (Type I and Type II), and about $11.99h_{cc}$ (1.78 nm) for neurofilaments and α-internexin (Type III and Type IV). Within the limitations of the data, these values are the same.

Little is known about the conformation of linker L12. In all cases except Type VI lamin molecules, it is predicted that the structure will not be α-helical. There is some similarity between chain types in their consensus sequences, and it follows that common structural features in L12 are likely across all chain types. The variations in sequence length could be taken up as external loops. A portion of the sequences do exhibit a conserved character in which an apolar-X motif is repeated four times (North *et al.*, 1994). It is possible that this will form a short length of β-structure. Noting that there are two chains in the molecule and that antiparallel molecules overlap their L12 segments to a large degree, a four-chain β-sheet or equivalent structure could occur. The evidence as yet is weak and further experimental data will be required before this idea will gain significant credence.

Electron microscope observations on individual IF molecules have shown that the molecules can bend, and even fold, over the L12 segment (Steven *et al.*, 1989). This is consistent with its predicted flexibility (Conway *et al.*, 1989; Smith *et al.*, 2002).

Gill *et al.* (1990) have engineered additional residues into linker L12. The result of these studies was that relatively normal filament formation was able to proceed. This indicates (but does not prove) that additional residues, which allow flexibility to be maintained, can be tolerated without compromising filament formation. However, mutations in which flexibility is lost or decreased appear to lead directly to malformation of IFs. The flexibility index for L12 is less than for linker L1, but it is clearly still an important feature of the structure.

F. Segment 2A

Segment 2A is conserved in length (19 residues) across all chain types. It has an underlying heptad repeat characteristic of a coiled-coil structure, as well as an acidic and basic residue period (about 9.8 residues long) first reported by Crewther *et al.* (1983). This is less easily seen by considering segment 2A alone (since it is so short), but is readily observed when segment 2A is considered as an extension to segment 2B (see Section II.H). Within segment 2A, there is a conserved basic residue in position 10 and a conserved acidic residue in position 14, thus allowing the formation of a conserved and potentially stabilizing intrahelical ionic interaction of the type i to $i + 4$ (Smith *et al.*, 2002).

The mode of molecular assembly of K5/K14 epidermal keratin has been studied in detail by making point mutations of charged residues along the chains, and then studying the aggregation of equimolar mixtures of wild type and mutant chains (Mehrani *et al.*, 2001). Using this technique, it has been possible to show that the conserved residue D1 and the less well-conserved D/N3 in segment 2A, and the conserved residue K31 in segment 1A, are all important for stability. These residues lie in close axial proximity in the A_{11} mode, and it is speculated that they are involved in making stabilizing intermolecular ionic interactions and hydrogen bonds.

G. Linker L2

Linker L2 is eight residues long for all IF chains, except for nestin (Type VI) where it is absent (Lendahl *et al.*, 1990). From crosslinking studies, the axially projected length can be determined (Parry and Steinert, 1999). For the Type I and Type II keratins, the projected length is about $4.77h_{cc}$ (0.71 nm). For Type III chains in vimentin IFs and Type IV chains in α-internexin (as well as their copolymerized assemblies), the projected lengths are about $7.79h_{cc}$ (1.16 nm). These data are limited and hence it is

likely that L2 will have a common projected length in all IF protein molecules of about 1 nm.

As reported for segments 1A and 2A, point mutations were made in charged residues in K5/K14 epidermal keratin. The stabilities of the resulting aggregates were measured to ascertain, in particular, some of the more important residues involved in stabilizing the A_{11} mode of molecular assembly. The results show that the conserved residues 4E and 6E in linker L2 were essential for stability, and that residue D7 was also important (Mehrani *et al.*, 2001). This group of residues lies immediately opposite to D/N9 and the conserved residue R10 in segment 1A. The residues were expected to interact favorably to provide the observed stability through both hydrogen bonds and intermolecular interactions.

L2 represents the region over which the coiled-coil conformation undergoes a significant azimuthal change between that present in segment 2A and that found in segment 2B. The underlying heptad repeat suffers a large phase shift at L2, equivalent to the insertion of five residues or deletion of two residues in an otherwise continuous heptad structure. The position of the apolar stripe on the surface of the individual α-helices thus changes from being internal to external. As this cannot occur *in vivo*, it follows that there must be a significant remodeling of the coiled coil to allow the appropriate phasing of the heptad repeat to be restored. In effect, the result is that there must be a marked azimuthal shift between the C-terminal end of segment 2A and the N-terminal end of segment 2B. There is a wide variety of possible conformations that are consistent with this idea and the actual amino acid sequences (see, for example, Conway and Parry, 1990; North *et al.*, 1994).

H. Segment 2B

The length of segment 2B is absolutely conserved (121 residues) in all IF chains, and over its entire range it exhibits a heptad substructure. However, there is a conserved stutter close to the center point of segment 2B and this breaks the phasing of the heptad repeat (Fig. 5). Brown *et al.* (1996) have categorized all heptad discontinuities and shown that these fall into three significant groups—stutters, stammers, and skips. They correspond to deletions of three, four, and six residues respectively. Conformationally, a stutter results in a local unwinding of the coiled coil such that the two strands lie approximately parallel to one another and to the axis of the molecule (Brown *et al.*, 1996; Strelkov and Burkhard, 2002). The perfect conservation of the stutter in all IF proteins points clearly to its fundamental importance. In effect, it will allow some fine tuning of the azimuthal coordinate of the coiled coil to occur, thus facilitating appropriate molecular

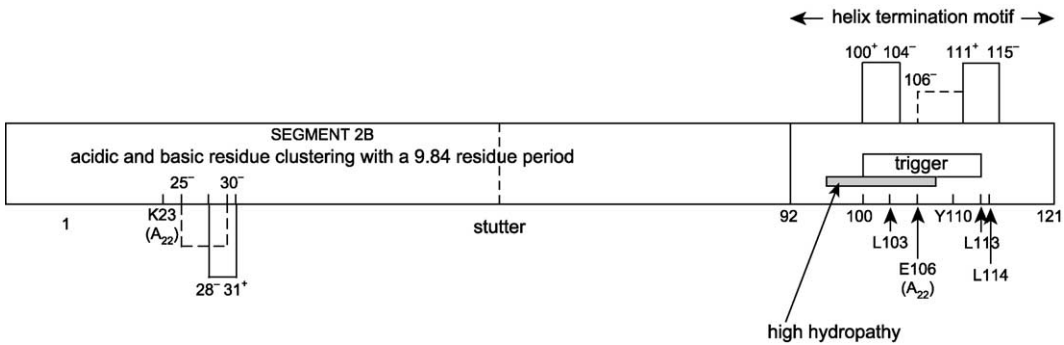


FIG. 5. Schematic diagram of segment 2B in intermediate filament chains showing generally conserved features. These include the highly conserved residues K23, L103, E106, Y110, L113, and L114; the conserved intra- and interchain ionic interactions that are represented by solid and dotted lines, respectively; the trigger motif (residues 100–113) that acts as a particularly stable region that ''nucleates'' coiled-coil formation; a region of particularly high hydropathy (residues 95–108); the helix termination motif (residues 92–121); a stutter in the heptad substructure at which point the coiled-coil structure unwinds locally. The entire segment displays a regular disposition of acidic and basic residues, each with a period of about 9.84 residues. These periods are approximately out of phase with one another.

interaction and assembly. Removal of the stutter in vimentin by inserting three residues to allow the heptad phasing to be continued uninterrupted has a negative effect on filament assembly, and only very short IF are formed *in vitro* (Herrmann and Aebi, 1998b). The conformational significance of the stutter is beyond question.

The C-terminal end of segment 2B contains a very highly conserved sequence of the form E–Y–Q–X–L–L–D/N–V–K–X–R/A–L–D/E–X–E–I–A–T–Y–R–K/R–L–L–E–G–E–E/D–X–R–L/N/I. This is known as the *helix termination motif* and encompasses residues 92–121. It can be regarded as the counterpart of the helix initiation motif described earlier in segment 1A (Fig, 5). The termination motif is important for other reasons, too. First, a trigger motif has been found in epidermal keratin (K5/K14) IFs, and this spans residues 100–113 (Wu *et al.*, 2000) within the helix termination motif. Its sequence is A–L–L–D–V–E–I–A–T–Y–R–K–L in K5 and T–R–L–E–Q–E–I–A–T–Y–R–R–L in K14. Second, all of the most highly conserved residues in segment 2B across all IF types occur in the helix termination motif (Smith *et al.*, 2002): the residues are L103, E106, Y110, L113, and L114. Third, the conserved intrachain ionic interactions of the i to $i + 4$ type are all in this same region: they are 100K-104D/E and 111R-115E (Smith *et al.*, 2002). Fourth, one of the two conserved interchain ionic interactions also lies in this region. This is between an acidic residue in position 106*g* and a basic residue in position 111*e*. The second is between an acidic residue in position 25*g* and a basic residue in position 30*e*. (In passing, it should be noted that the sole conserved i to $i + 3$ intrachain ionic interaction lies between 28E/D and 31R: Smith *et al.*, 2002). Fifth, the highest hydropathy calculated for IF chains occurs for residues 95–108 (Smith *et al.*, 2002). Sixth, in the A_{22} mode of molecular interaction, Wu *et al.* (2000) have shown that E106 is important for stability, as is K23 in the same segment. These residues lie in close axial alignment in the A_{22} mode and are likely to form an ionic interaction.

The helix termination motif may well be unique in containing so many special features of structural and functional importance in such a short length of sequence. Indeed, the experimental observation that the C-terminal end of segment 2B was a mutation hot spot that would lead to a large number of diseases (Parry and Steinert, 1999) was readily predictable on theoretical grounds. There are, as stated above, a multitude of key residues in this region that are involved in almost every aspect of the structure and assembly of IF molecules.

One aspect of the assembly of IF molecules that has received scant attention thus far relates to the small head-to-tail overlap that occurs between similarly directed molecules. This is defined by mode A_{CN}. Its value is generally about 7 h_{cc} (7.03 h_{cc} or 1.04 nm in epidermal keratin,

5.67 h_{cc} or 0.84 nm in reduced trichocyte keratin, 4.31 h_{cc} or 0.64 nm in vimentin, and 10 h_{cc} or 1.49 nm in α-internexin), although in oxidized trichocyte there is a head-to-tail gap of 8.43 h_{cc} or 1.25 nm (Parry, 1995; Parry *et al.,* 2001; Steinert *et al.,* 1993a,b,c, 1999a; Wang *et al.,* 2000). Structural data on this region of segment 2B reveal that the α-helical strands have separated over the last ten residues (Herrmann *et al.,* 2000). Knowing also that the strands of segment 1A can likewise be structurally independent of one another (Strelkov *et al.,* 2002), it is interesting to speculate that the overlap can, under assembly conditions, be conformationally akin to a four-α-helix motif with two similarly directed strands contributed by segment 1A and two oppositely directed strands contributed by segment 2B (Smith *et al.,* 2002). A comparable structure has been considered by Carolyn Cohen for the head-to-tail overlap of similarly directed tropomyosin molecules.

As in segment 1B, the linear distributions of the acidic and the basic residues are nonrandom. The statistically significant period determined by Fourier analysis (Parry and Fraser, 1985) is 9.84 h_{cc} (1.46 nm). An important point is that the difference in phasing between the acidic and the basic periods is close, but not equal, to 180 degrees. Thus, an appropriate family of axial staggers between two 2B segments (parallel or antiparallel) will necessarily facilitate intermolecular ionic interactions. However, antiparallel alignments will be favored over parallel ones because more interactions are possible. This arises directly from the difference in phasing of the identical acidic and basic periods not being equal to 180 degrees. The A_{22} mode of interaction is largely specified by these types of interactions though the contributions of hydrogen bonds, and apolar interactions are also likely to be important.

I. Tail Domain

1. Epidermal Keratins (Type I/Type II)

The epidermal keratins have a tail subdomain structure closely parallel to that described for the heads in Section II.A.1. Immediately C-terminal to the rod domain exists a stretch of highly conserved sequence across Type II chains (Steinert *et al.,* 1985). This is termed the *H2 subdomain* and is 20 residues long. An equivalent conserved region in Type I chains, however, is totally absent. C-terminal to H2 lies a sequence that varies between chains within a particular type, but which is rich in glycine and serine residues. This V2 subdomain, present in both Type I and Type II chains, has a length that lies in the range 0–110 and 25–125 residues in the two chain types, respectively. Human K1 and K10 chains both show

polymorphism in their V2, but not V1, subdomains (Fig. 2b). Insertions and deletions are more prevalent in the Type II chains, but have also been recognized in the Type I chains (Korge *et al.*, 1992a,b; Mischke, 1993; Parry and Steinert, 1999). The presence of a glycine loop conformation, of course, readily allows such polymorphism to occur without major structural implications.

2. *Trichocyte Keratins (Type I/Type II)*

The tail domain in trichocyte Type I chains contains a tenfold P–C–X repeat, seven of them contiguously (Parry and North, 1998). Studies of its likely conformation suggest that it will most likely adopt a left-handed polyglycine II structure with three residues per turn (Fig. 3a). The cysteine residues would thus lie along one edge of the structure, where they would be in positions to form disulphide bonds with cysteine residues in a similar structure from a different molecule. The role of this sequence motif would appear to be that of stabilizing molecular assembly within the IFs.

The Type II trichocyte chain, however, contains a quite different structural motif—a right-handed, four-stranded, twisted antiparallel β-sheet (Fig. 3b). The tetrapeptides linking strands 1–2 and strands 2–3 are homologous and have sequences of S-S-S-R and S-G-S-R/A. Similarly, the tetrapeptides linking strands 3–4 and strand 4 with the remainder of the tail region have homologous sequences of A–P–C–S and A–P–C–G, respectively. In all four cases, the sequences are predicted to form a β-turn structure. The β-strands are characterized by a two-residue apolar repeat, resulting in one face of the β-sheet being almost totally apolar. The other face would contain the cysteine residues. This structure immediately suggests a role *in vivo* whereby the apolar face in the β-sheet of one molecule would interact with the same domain in a second molecule in the IFs. This would again lead to stabilization of the molecular aggregate within the filament.

3. *Desmin, Vimentin, Glial Fibrillary Acidic Protein, and Peripherin (Type III), Neurofilaments and α-Internexin (Type IV), Lamins (Type V), and Nestin (Type VI)*

Almost nothing is known of the role played by individual amino acid residues in the tail domains of Type III IF proteins. Also, very little sequence regularity has been observed. Nonetheless, the degree of sequence similarity in the tail domains of these proteins suggests that a role analogous to that played by the H subdomains in the keratins is not

unlikely. In other words, the tails may be important in stabilizing the assembly of Type III IF proteins at the oligomeric stage, possibly at the two- to four-molecule level.

In contrast to the lack of obvious subdomain structure in the Type III IF proteins, the organization of the tail domains in the Type IV IF proteins is well established, and consists of a number of regions of special sequence and character (Fig. 6; Parry and Steinert, 1995). These include the glutamic acid rich region (E segment), the conserved lysine-glutamic acid region (KE segment), the lysine-serine-proline repeats (KSP segment), and the terminal lysine-glutamic acid-proline segments (KEP segment). The KSP repeats are the sites of phosphorylation, but the extent of these differs markedly between species (see, for example, the work of Napolitano *et al.,* 1987 and Chin and Liem, 1990). The role of the E segment is believed to be important in IF assembly (Birkenberger and Ip, 1990).

The tail domain of nestin contains 22-residue (human) and 44-residue repeats (hamster and rat) (Parry and Steinert, 1999; Steinert *et al.,* 1999b). The 44-residue repeat of hamster is E–E–D–Q–L/R–V–E–R/T–L–I–E–K–E– G–Q–E–S–L–S–S–P–E and E–E–D–Q–E–T–D–R–P–L–E–K–E–N–G–E–P–L–K–P–V–E and, as divided here, can easily be seen to consist of a pair of related but nonidentical 22-residue sequences. Interestingly, even these 22-residue repeats are quasi-halved, indicating an underlying 11-residue structure in all of the species studied. The secondary conformation is unknown, but there is an expectation of high α–helix content.

4. Invertebrate IF

Uniquely, the cephalocordate C1 and C2 chains both contain heptad repeats of overall length 64 residues in their tail domains (Riemer *et al.,* 1998). There is no direct evidence, however, of the role that these repeats play *in vivo*, but it is evident that such sequences are not present by chance. There can be no doubt that this repeat is designed to facilitate oligomerization between chains.

J. Features in Head and Tail Domains

The head and tail domains in keratin molecules generally contain a multitude of sites that allow keratin IFs to form covalent bonds with other proteins. For example, in the case of the keratin IFs in the inner root sheath, lysine to glutamine crosslinks have been characterized between the head and (more often) the tail domains in (generally) both Type I and Type II chains, envoplakin, epiplakin, the SPR2 (small proline-rich) family, and trichohyalin (Steinert *et al.,* 2003). Likewise, in noninner root

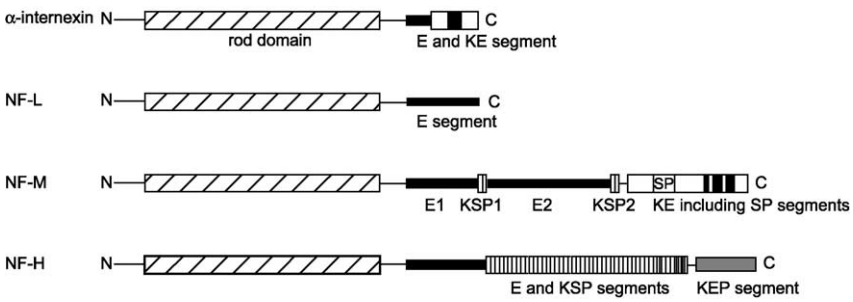


FIG. 6. Schematic representation of the tail domains in Type IV intermediate filament chains. The rod domain, indicated by a shaded rectangle, does not show its known substructure. The four chains illustrated here are, from the top downwards, α-internexin, and the low (NF-L), medium (NF-M), and high molecular weight chains (NF-H) from neurofilaments. The characters of the segments are indicated by the one letter code. Figure is redrawn from Parry and Steinert (1995) and is based on an original by Shaw.

sheath epidermal keratin chains, crosslinks have been identified between the lysine residue in the K-S-I-S-I-S motif in the V1 subdomain of the Type II head domain and glutamine residues in cystatin, elafin, envoplakin, involucrin, loricrin, and the SPR1, 2, and 3 families of the cornified cell envelope (Candi *et al.*, 1998; Steinert and Marekov, 1995, 1997). It is interesting and pertinent that the lower crosslink density in these IFs probably arises in large part from the fact that only a single contributory lysine site occurs, and in just one of the two chain types. In both of the examples listed here, the process is mediated by transglutaminase crosslinking.

Parry and Steinert (1999) noted that posttranslational modifications were found almost (but not quite) exclusively in the head and tail domains of all IF proteins. Nothing in the five years since their review has changed that conclusion, though more examples of amino acid modification are now known. A selection of posttranslational modifications is noted below, but a comprehensive list has not been attempted and is considered beyond the scope of this review. Useful references to posttranslational modifications include those of Quinlan *et al.* (1994), Parry and Steinert (1999), and Herrmann and Aebi (2000).

A posttranslational modification of major significance in IFs relates to the phosphorylation and dephosphorylation events that occur at serine and/or threonine residues within specific recognition sites. These sites are located entirely in the head and tail domains. It has now been widely established that the action of kinases which lead to phosphorylation, and

phosphotases which cause dephosphorylation, are the crucial biochemical actions that regulate (or partially regulate) the assembly and the turnover of a wide variety of IFs *in vivo*. They also can modify the interactions between IFs and various cellular entities. For instance, during cell division the phosphorylation of some IFs (cytokeratins, vimentin, lamin) results in filament disassembly; the inverse event—dephosphorylation—results in spontaneous self-assembly to form intact filaments.

In the medium and large neurofilament chains, where there are numerous K-S-P phosphorylation sites in the tail domains, the effect of phosphorylation is quite different. It has no visible effect on the state of filament assembly. It does, however, appear to be particularly important in determining axonal diameter (and concomitant conduction velocity), as well as transport properties and association with other cytoskeletal components. Experimentally, numerous phosphorylation sites have been shown to exist in a wide variety of IF proteins. Many others have been proposed on the basis of sequence motifs consistent with sites of known kinases. It has also been shown that mutations in which phosphorylation sites have been changed (see, for example, S35A in keratin 19) lead to various pathologies, including malformations in the filament assembly.

Other posttranslational modifications seen in IFs include proteolysis (cytokeratin, lamin, neurofilaments: Klymkowsky *et al.,* 1991; Nigg, 1992) and, in the case of lamin, an isoprenylation event at the CaaX box. The latter, which is unique among IF chains, occurs at the C-terminal end of the chain (Hennekes and Nigg, 1994; Nigg, 1992) and has the effect of targeting lamin chains to the nuclear envelope. Another quite different posttranslational modification is that of mono-ADP-ribosylation, which appears to modify desmin assembly (Yuan *et al.,* 1999). Glycosylation is yet another posttranslational modification, and a wide variety of IF chains have been reported as being glycosylated. These include some keratins, lamins, and some neurofilament chains. Glycosylation sites, like those of many other (but not all) posttranslational modifications, occur in both the head and tail domains (Quinlan *et al.,* 1994).

III. Summary

The quarter-century that elapsed since the first sequences of IF proteins were published and analyzed has seen a great deal of new information about the chain and molecular structure, as well as the modes by which the molecules assemble to form viable filaments. As the research has progressed, it has become possible to identify the structural and/or functional role of various subdomains within the IF chains, and even of individual

amino acid residues. This review has attempted to categorize that knowledge, to identify residues that are highly conserved and which are of particular structural/functional importance, and to ascertain both intra- and interchain interactions that are especially crucial in either stabilizing or nucleating the coiled-coil structure.

It is evident that specific residue-related features observed in the rod domain of one particular IF chain type are frequently not observed in the same place (if at all) in the other chain types and, of course, the large differences in head and tail structure and sequence between chain types preclude identical roles from occurring there as well. The highly specialized sequence characteristics thus define unique structural assemblies and functions, while still maintaining a high degree of structural uniformity, particularly in the manner of rod domain assembly. Even in this case, small but important differences in the A_{11}, A_{22}, and A_{12} modes occur. Distinct IF structures have been identified for: (1) unoxidized trichocyte keratin and epidermal keratin; (2) oxidized trichocyte keratin; (3) Type III and IV IF proteins; and (4) the nuclear lamins.

Much more is required before we will have a detailed understanding of IF structure, but real progress has been made. We can realistically expect that the next ten years will see new studies completed that will add greatly to our knowledge. Crystal studies of IF protein fragments, now in progress, will add immeasurably to that understanding.

References

Bilak, S. R., Sernett, S. W., Bilak, M. M., Bellin, R. M., Stromer, M. H., Huiatt, T. W., and Robson, R. M. (1998). Properties of the novel intermediate filament protein synemin and its identification in mammalian muscle. *Archiv. Biochem. Biophys.* **355,** 63–76.

Birkenberger, L., and Ip, W. (1990). Properties of the desmin tail domain: Studies using synthetic peptides and antipeptide antibodies. *J. Cell Biol.* **111,** 2063–2075.

Brown, J. H., Cohen, C., and Parry, D. A. D. (1996). Heptad breaks in α-helical coiled coils: Stutters and stammers. *Proteins* **26,** 134–145.

Burkhard, P., Kammerer, R. A., Steinmetz, M. O., Bourenkov, G. P., and Aebi, U. (2000). The coiled-coil trigger site of the rod domain of cortexillin I unveils a distinct network of interhelical and intrahelical salt bridges. *Struct. Fold. Des.* **8,** 223–230.

Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001). Coiled coils: A highly versatile protein folding motif. *Trends Cell Biol.* **11,** 82–88.

Candi, E., Tarcsa, E., DiGiovanna, J. J., Compton, J. G., Elias, P. M., Marekov, L. N., and Steinert, P. M. (1998). A highly conserved lysine residue on the head domain of Type II keratins is essential for the attachment of keratin intermediate filaments to the cornified cell envelope through isopeptide crosslinking transglutaminases. *Proc. Natl. Acad. Sci. USA* **95,** 2067–2072.

Chin, S. S. M., and Liem, R. K. H. (1990). Transfected rat high-molecular-weight neurofilament (NF-H) coassembles with vimentin in a predominantly nonphosphorylated form. *J. Neurosci.* **10,** 3714–3726.

Ching, G. Y., and Liem, R. K. (1993). Assembly of Type IV neuronal intermediate filaments in nonneuronal cells in the absence of preexisting cytoplasmic intermediate filaments. *J. Cell Biol.* **122,** 1323–1335.

Conway, J. F., and Parry, D. A. D. (1990). Structural features in the heptad substructure and longer range repeats of two-stranded α-fibrous proteins. *Int. J. Biol. Macromol.* **12,** 328–334.

Conway, J. F., Fraser, R. D. B., MacRae, T. P., and Parry, D. A. D. (1989). Protein Chains in Wool and Epidermal Keratin IF: Structural features and spatial arrangement. *In* ''The Biology of Wool and Hair'' (G. E. Rogers, P. J. Reis, K. A. Ward, and R. C. Marshall, Eds.), pp. 127–144. Chapman and Hall, London and New York.

Crewther, W. G., Dowling, L. M., Steinert, P. M., and Parry, D. A. D. (1983). Structure of intermediate filaments. *Int. J. Biol. Macromol.* **5,** 267–274.

Downing, S. W., Spitzer, R. H., Salo, W. L., Downing, J. S., Saidel, L. J., and Koch, E. A. (1981). Threads in the hagfish slime gland thread cells: Organization, biochemical features and length. *Science* **212,** 326–328.

Fisher, D. Z., Chaudhary, N., and Blobel, G. (1986). cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. *Proc. Natl. Acad. Sci. USA* **83,** 6450–6454.

Fraser, R. D. B., MacRae, T. P., Suzuki, E., and Parry, D. A. D. (1985). Intermediate filament structure. 2. Molecular interactions in the filament. *Int. J. Biol. Macromol.* **7,** 258–274.

Fraser, R. D. B., MacRae, T. P., Parry, D. A. D., and Suzuki, E. (1986). Intermediate filaments in α-keratins. *Proc. Natl. Acad. Sci. USA* **83,** 1179–1183.

Gill, S. R., Wong, P. C., Monteiro, M. J., and Cleveland, D. W. (1990). Assembly properties of dominant and recessive mutations in the small mouse neurofilament (NF-L) subunit. *J. Cell Biol.* **111,** 2005–2019.

Goulielmos, G., Gounari, F., Remington, S., Müller, S., Häner, M., Aebi, U., and Georgatos, S. D. (1996). Filensin and phakinin form a novel type of beaded intermediate filaments and coassemble *de novo* in cultured cells. *J. Cell Biol.* **132,** 643–655.

Hatzfeld, M., and Burba, M. (1994). Function of Type I and Type II keratin head domains: Their role in dimer, tetramer and filament formation. *J. Cell Sci.* **107,** 1959–1972.

Hatzfeld, M., and Weber, K. (1990). The coiled-coil of *in vitro* assembled keratin filaments is a heterodimer of Type I and II keratins: Use of site-specific mutagenesis and recombinant protein expression. *J. Cell Biol.* **110,** 1199–1210.

Hennekes, H., and Nigg, E. A. (1994). The role of isoprenylation in membrane attachment of nuclear lamins: A single point mutation prevents proteolytic cleavage of the lamin A precursor and confers membrane binding properties. *J. Cell Sci.* **107,** 1019–1029.

Herrling, J., and Sparrow, L. G. (1991). Interactions of intermediate filament proteins from wool. *Int. J. Biol. Macromol.* **13,** 115–119.

Herrmann, H., and Aebi, U. (1998a). Intermediate filament assembly: Fibrillogenesis is driven by decisive dimer-dimer interactions. *Curr. Opin. Struct. Biol.* **8,** 177–185.

Herrmann, H., and Aebi, U. (1998b). Structure, Assembly and Dynamics of Intermediate Filaments. *In* ''Subcellular Biochemistry: Intermediate Filaments'' (H. Herrmann and J. R. Harris, Eds.), Vol 31. pp. 319–362. Plenum Press, New York.

Herrmann, H., and Aebi, U. (2000). Intermediate filaments and their associates: Multi-talented structural elements specifying cyto-architecture and cyto-dynamics. *Curr. Opin. Cell Biol.* **12,** 79–90.

Herrmann, H., Hofmann, I., and Franke, W. W. (1992). Identification of a nonapeptide motif in the vimentin head domain involved in intermediate filament assembly. *J. Mol. Biol.* **223,** 637–650.

Herrmann, H., Häner, M., Brettel, M., Ku, N-O., and Aebi, U. (1999). Characterization of distinct early assembly units of different intermediate filament proteins. *J. Mol. Biol.* **286,** 1403–1420.

Herrmann, H., Strelkov, S. V., Feja, B., Rogers, K. R., Brettel, M., Lustig, A., Häner, M., Parry, D. A. D., Steinert, P. M., Burkhard, P., and Aebi, U. (2000). The intermediate filament protein consensus motif of helix 2B: Its atomic structure and contribution to assembly. *J. Mol. Biol.* **298,** 817–832.

Herrmann, H., Hesse, M., Reichenzeller, M., Aebi, U., and Magin, T. M. (2003). Functional complexity of intermediate filament cytoskeletons: From structure to assembly to gene ablation. *Int. Rev. Cytol.* **223,** 83–175.

Klymkowsky, M. W., Maynell, L. A., and Nislow, C. (1991). Cytokeratin phosphorylation, cytokeratin filament severing and the solubilization of the maternal mRNA *Vg1*. *J. Cell Biol.* **114,** 787–797.

Koch, E. A., Spitzer, R. H., Pithawalla, R. B., and Downing, S. W. (1991). Keratin-like components of gland thread cells modulate the properties of mucus from hagfish (Eptatretus stouti). *Cell Tiss. Res.* **264,** 79–86.

Koch, E. A., Spitzer, R. H., Pithawalla, R. B., and Parry, D. A. D. (1994). An unusual intermediate filament subunit from the cytoskeletal biopolymer released extracellularly into seawater by the primitive hagfish (Eptatretus stouti). *J. Cell Sci.* **107,** 3133–3144.

Koch, E. A., Spitzer, R. H., Pithawalla, R. B., Castillos, F. A., and Parry, D. A. D. (1995). Hagfish biopolymer: A Type I/Type II homologue of epidermal keratin intermediate filaments. *Int. J. Biol. Macromol.* **17,** 283–292.

Korge, B. P., Gan, S.-Q., McBride, O. W., Mischke, D., and Steinert, P. M. (1992a). Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops. *Proc. Natl. Acad. Sci. USA* **89,** 910–914.

Korge, B. P., Compton, J. G., Steinert, P. M., and Mischke, D. (1992b). The two size alleles of human keratin 1 are due to a deletion in the glycine-rich carboxyl-terminal V2 subdomain. *J. Invest. Dermatol.* **99,** 697–702.

Kühnel, K., Jarchau, T., Wolf, E., Schlichting, I., Walter, U., Wittinghofer, A., and Strelkov, S. V. (2004). The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proc. Natl. Acad. Sci. USA* **101,** 17027–17032.

Lee, M-K., Xu, Z., Wong, P-C., and Cleveland, D. W. (1993). Neurofilaments are obligate heteropolymers *in vivo*. *J. Cell Biol.* **122,** 1337–1350.

Lendahl, U., Zimmerman, L. B., and McKay, R. D. G. (1990). CNS stem cells express a new class of intermediate filament protein. *Cell* **60,** 585–595.

Mack, J. W., Torchia, D. A., and Steinert, P. M. (1988). Solid-state NMR studies of the dynamics and structure of mouse keratin intermediate filaments. *Biochemistry* **27,** 5418–5426.

McKeon, F. D., Kirschner, M. W., and Caput, D. (1986). Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature (London)* **319,** 463–468.

McLachlan, A. D., and Stewart, M. (1982). Periodic charge distribution in the intermediate filament proteins desmin and vimentin. *J. Mol. Biol.* **162,** 693–698.

Mehrani, T., Wu, K. C., Morasso, M. I., Bryan, J. T., Marekov, L. N., Parry, D. A. D., and Steinert, P. M. (2001). Residues in the 1A rod domain segment and linker L2 are required for stabilizing the A_{11} molecular alignment mode in keratin intermediate filaments. *J. Biol. Chem.* **276,** 2088–2097.

Mischke, D. (1993). Frequencies of human keratin 10 alleles. *Hum. Mol. Genet.* **2,** 618.

Mücke, N., Wedig, T., Bürer, A., Marekov, L. N., Steinert, P. M., Langowski, J., Aebi, U., and Herrmann, H. (2004). Molecular and biophysical characterization of assembly-starter units of human vimentin. *J. Mol. Biol.* **340,** 97–114.

Napolitano, E. W., Chin, S. S. M., Colman, D. R., and Liem, R. K. H. (1987). Complete amino acid sequence and *in vitro* expression of rat NF-M, the middle molecular weight neurofilament protein. *J. Neurosci.* **7,** 2590–2599.

Nigg, E. A. (1992). Assembly-disassembly of the nuclear lamina. *Curr. Opin. Cell Biol.* **4,** 105–109.

North, A. C. T., Steinert, P. M., and Parry, D. A. D. (1994). Coiled-coil stutter and link segments in keratin and other intermediate filaments: A computer modeling study. *Proteins* **20,** 174–184.

Özbek, S., Müller, J. F., Figgemeier, E., and Stetefeld, J. (2004). Insights into thermostability at atomic level: The crystal structure of the right-handed tetrabrachion coiled coil of Staphylothermus marinus at 1.4Å resolution. *J. Mol. Biol.* (in press).

Parry, D. A. D. (1995). Hard α-keratin IF: A structural model lacking a head-to-tail molecular overlap but having hybrid features characteristic of both epidermal keratin and vimentin IF. *Proteins* **22,** 267–272.

Parry, D. A. D. (1997). Protein chains in hair and epidermal keratin IF: Structural features and spatial arrangements. *In* ''Formation and Structure of Human Hair'' (P. Jollés, H. Zahn, and H. Höcker, Eds.), pp. 177–207. Birkhäuser, Basel.

Parry, D. A. D., and Fraser, R. D. B. (1985). Intermediate filament structure I. Analysis of IF protein sequence data. *Int. J. Biol. Macromol.* **7,** 203–213.

Parry, D. A. D., and North, A. C. T. (1998). Hard α-keratin intermediate filament chains: Substructure of the N- and C-terminal domains and the predicted structure and function of the C-terminal domains of Type I and Type II chains. *J. Struct. Biol.* **122,** 67–75.

Parry, D. A. D., Steinert, P. M., (1995). ''Intermediate filament structure,'' pp. 1–183. Springer-Verlag, Heidelberg.

Parry, D. A. D., and Steinert, P. M. (1999). Intermediate filaments: Molecular architecture, assembly, dynamics and polymorphism. *Quart. Rev. Biophys.* **32,** 99–187.

Parry, D. A. D., Crewther, W. G., Fraser, R. D. B., and MacRae, T. P. (1977). Structure of α-keratin: Structural implication of the amino acid sequences of the Type I and Type II chain segments. *J. Mol. Biol.* **113,** 449–454.

Parry, D. A. D., Steven, A. C., and Steinert, P. M. (1985). The coiled-coil molecules of intermediate filaments consist of two parallel chains in exact axial register. *Biochem. Biophys. Res. Commun.* **127,** 1012–1018.

Parry, D. A. D., Conway, J. F., and Steinert, P. M. (1986). Structural studies on lamin: Similarities and differences between lamin and intermediate-filament proteins. *Biochem. J.* **238,** 305–308.

Parry, D. A. D., Marekov, L. N., and Steinert, P. M. (2001). Subfilamentous protofibril structures in fibrous proteins: Cross-linking evidence for protofibrils in intermediate filaments. *J. Biol. Chem.* **276,** 39253–39258.

Parry, D. A. D., Marekov, L. N., Steinert, P. M., and Smith, T. A. (2002). A role for the 1A and L1 rod domain segments in head domain organization and function of

intermediate filaments: Structural analysis of trichocyte keratin. *J. Struct. Biol.* **137,** 97–108.

Quinlan, R., Hutchison, C., and Lane, B. (1994). Intermediate filament proteins. *Protein Profile* **1,** 779–911.

Ralton, J. E., Lu, X., Hutcheson, A. M., and Quinlan, R. A. (1994). Identification of two N-terminal non-α helical domain motifs important in the assembly of glial fibrillary acidic protein. *J. Cell Sci.* **107,** 1935–1948.

Riemer, D., Karabinos, A., and Weber, K. (1998). Analysis of eight cDNAs and six genes for intermediate filament (IF) proteins in the Cephalochordate *Branchiostoma* reveals differences in the IF multigene families of lower chordates and the vertebrates. *Gene* **211,** 361–373.

Shaw, G. (1998). ''Neurofilaments,'' pp. 1–287. Springer-Verlag, Berlin.

Smith, T. A., Strelkov, S. V., Burkhard, P., Aebi, U., and Parry, D. A. D. (2002). Sequence comparisons of intermediate filament chains: Evidence of a unique functional/structural role for coiled-coil segment 1A and Linker L1. *J. Struct. Biol.* **137,** 128–145.

Smith, T. A., Hempstead, P. D., Palliser, C. C., and Parry, D. A. D. (2003). Modeling α-helical coiled-coil interactions: The axial and azimuthal alignment of 1B segments from vimentin intermediate filaments. *Proteins* **50,** 207–212.

Smith, T. A., Steinert, P. M., and Parry, D. A. D. (2004). Modeling effects of mutations in coiled-coil structures: Case study using Epidermolysis Bullosa Simplex mutations in segment 1A of K5/K14 intermediate filaments. *Proteins* **55,** 1043–1052.

Steinert, P. M. (1990). The two-chain coiled-coil molecule of native epidermal keratin intermediate filaments is a Type I-Type II heterodimer. *J. Biol. Chem.* **265,** 8766–8774.

Steinert, P. M., and Marekov, L. N. (1995). The Proteins elafin, filaggrin, keratin intermediate filaments, loricrin and small proline-rich proteins are isodipeptide cross-linked components of the human epidermal cornified cell envelope. *J. Biol. Chem.* **270,** 17702–17711.

Steinert, P. M., and Marekov, L. N. (1997). Direct evidence that involucrin is a major early isopeptide crosslinked component of the keratinocyte cornified cell envelope. *J. Biol. Chem.* **272,** 2021–2030.

Steinert, P. M., and Parry, D. A. D. (1993). The conserved H1 domain of the Type II keratin 1 chain plays an essential role in the alignment of nearest neighbor molecules in mouse and human keratin 1/keratin 10 intermediate filaments at the two- to four-molecule level of structure. *J. Biol. Chem.* **268,** 2878–2887.

Steinert, P. M., Parry, D. A. D., Idler, W. W., Johnson, L. D., Steven, A. C., and Roop, D. R. (1985). Amino acid sequences of mouse and human epidermal Type II keratins of M_r 67000 provide a systematic basis for the structural and functional diversity of the end domains of keratin intermediate filament subunits. *J. Biol. Chem.* **260,** 7142–7149.

Steinert, P. M., Mack, J. W., Korge, B. P., Gan, S-Q., Haynes, S. R., and Steven, A. C. (1991). Glycine loops in proteins: Their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. *Int. J. Biol. Macromol.* **13,** 130–139.

Steinert, P. M., Marekov, L. N., Fraser, R. D. B., and Parry, D. A. D. (1993a). Keratin intermediate filament structure: Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J. Mol. Biol.* **230,** 436–452.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1993b). Conservation of the structure of keratin intermediate filaments: Molecular mechanism by which

different keratin molecules integrate into pre-existing keratin intermediate filaments during differentiation. *Biochemistry* **32,** 10046–10056.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1993c). Diversity of intermediate filament structure: Evidence that the alignment of coiled-coil molecules in vimentin is different from that in keratin intermediate filaments. *J. Biol. Chem.* **268,** 24916–24925.

Steinert, P. M., Marekov, L. N., and Parry, D. A. D. (1999a). Molecular parameters of Type IV α-internexin and Type IV-Type III α-internexin-vimentin copolymer intermediate filaments. *J. Biol. Chem.* **274,** 1657–1666.

Steinert, P. M., Chou, Y-H., Prahlad, V., Parry, D. A. D., Marekov, L. N., Wu, K. C., Jang, S-I., and Goldman, R. D. (1999b). A high molecular weight intermediate filament-associated protein in BHK-21 cells is nestin, a Type VI intermediate filament protein: Limited co-assembly *in vitro* to form heteropolymers with Type III vimentin and Type IV α-internexin. *J. Biol. Chem.* **274,** 9881–9890.

Steinert, P. M., Parry, D. A. D., and Marekov, L. N. (2003). Trichohyalin mechanically strengthens the hair follicle: Multiple cross-bridging roles in the inner root sheath. *J. Biol. Chem.* **278,** 41409–41419.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R. A. (2000). Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–776.

Steven, A. C., Mack, J. W., Trus, B. L., Bisher, M. E., and Steinert, P. M. (1989). Structure and assembly of intermediate filaments: Multi-faceted, myosin-like (but non-motile) cytoskeletal Polymers. *In* ''Cytoskeletal and Extracellular Proteins'' (U. Aebi and J. Engel, Eds.), pp. 15–26. Springer-Verlag, Berlin.

Strelkov, S. V., and Burkhard, P. (2002). Analysis of α-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *J. Struct. Biol.* **137,** 54–64.

Strelkov, S. V., Herrmann, H., Geisler, N., Lustig, A., Ivaninskii, S., Zimbelmann, R., Burkhard, P., and Aebi, U. (2001). Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *J. Mol. Biol.* **306,** 773–781.

Strelkov, S. V., Herrmann, H., Geisler, N., Wedig, T., Zimbelmann, R., Aebi, U., and Burkhard, P. (2002). Conserved segments 1A and 2B of the intermediate filament dimer: Their atomic structures and role in filament assembly. *EMBO J.* **21,** 1255–1266.

Suzuki, E., Crewther, W. G., Fraser, R. D. B., MacRae, T. P., and McKern, N. M. (1973). X-ray diffraction and infrared studies of an α-helical fragment from α-keratin. *J. Mol. Biol.* **73,** 275–278.

Wang, H., Parry, D. A. D., Jones, L. N., Idler, W. W., Marekov, L. N., and Steinert, P. M. (2000). *In vitro* assembly and structure of trichocyte keratin intermediate filaments: A novel role for stabilization by disulfide bonding. *J. Cell Biol.* **151,** 1459–1468.

Wu, K. C., Bryan, J. T., Morasso, M. I., Jang, S. I., Lee, J. H., Yang, J-M., Marekov, L. N., Parry, D. A. D., and Steinert, P. M. (2000). Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. *Mol. Biol. Cell* **11,** 3539–3558.

Yuan, J., Huiatt, T. W., Liao, C. X., Robson, R. M., and Graves, D. J. (1999). The effects of mono-ADP-ribosylation on desmin assembly-disassembly: Determination of the chemical features of bound ADP-ribose that prevent desmin filament formation. *Arch. Biochem. Biophys.* **363,** 314–322.

INTERMEDIATE FILAMENT ASSOCIATED PROTEINS

By KATHLEEN J. GREEN,* MICHAEL BÖHRINGER,† TODD GOCKEN,* AND JONATHAN C. R. JONES†

*Departments of Pathology and Dermatology and R.H. Lurie Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611;
†Department of Cell and Molecular Biology and R.H. Lurie Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611

I. Introduction 144
II. Plakin Family 146
A. Desmoplakin 149
B. BP230 (BPAG1) and its Isoforms 152
C. Plectin 154
D. Envoplakin and Periplakin 157
E. Epiplakin 159
III. Cell Surface–IF Interactions 159
A. Proteins in Desmosomes 159
B. Polycystin-1 161
C. Association of IFs with the Actin-Rich Cortex 162
IV. Other Crosslinkers not in the Plakin Family 168
A. Keratinocyte Crosslinkers and Bundlers 168
B. Proteins that Coordinate with Actin 171
V. IFAPs as Participants in a Signaling Scaffold and Cellular Metabolism 173
A. Stress Proteins/Chaperones 173
B. Proteins that Regulate Apoptosis 174
C. Interaction of IFs with Kinases and Cell Cycle Machinery 175
D. 14-3-3 Proteins 176
E. Interaction with Lipids and Membranes: Involvement in Membrane Trafficking 177
VI. Motors that Move IFs 178
VII. The Fine Line between IFs and IFAPs 180
VIII. Conclusions and Future Directions 181
References 182

Abstract

Intermediate filament associated proteins (IFAPs) coordinate interactions between intermediate filaments (IFs) and other cytoskeletal elements and organelles, including membrane-associated junctions such as desmosomes and hemidesmosomes in epithelial cells, costameres in striated muscle, and intercalated discs in cardiac muscle. IFAPs thus serve as critical connecting links in the IF scaffolding that organizes the cytoplasm and confers mechanical stability to cells and tissues. However, in recent years it has become apparent that IFAPs are not limited to structural

DOI: 10.1016/S0065-3233(04)70006-6
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

crosslinkers and bundlers but also include chaperones, enzymes, adapters, and receptors. IF networks can therefore be considered scaffolding upon which associated proteins are organized and regulated to control metabolic activities and maintain cell homeostasis.

I. Introduction

Intermediate filaments (IFs), along with actin-containing microfilaments and microtubules, comprise the three major nonmuscle cell cytoskeleton proteins. IFs fall into five major families that exhibit tissue specific expression patterns (Coulombe and Omary, 2002; Coulombe and Wong, 2004) (also see Chapter 5). Within the cell, IFs are associated with the nuclear surface and extend out into the cytoplasm, where they interact with other cytoskeletal tracks and cytoplasmic organelles such as mitochondria. They also impinge on the plasma membrane in certain tissues where they are tethered by ultrastructurally visible organelles and specialized membrane densities such as desmosomes and hemidesmosomes in epithelial cells, costameres in striated muscle, and intercalated discs in cardiac muscle. IFs thus serve as integrators of cytoplasmic space within cells and of cells within tissues, consequently conferring mechanical stability to organs.

Interactions between IFs and other cytoskeletal elements and organelles are coordinated by intermediate filament associated proteins (IFAPs); many of these proteins have emerged as critical regulators of tissue integrity in their own right (Coulombe and Wong, 2004; Fuchs and Yang, 1999; Herrmann and Aebi, 2000; Houseweart and Cleveland, 1999). In addition, however, it is clear that IFAPs are not limited to crosslinkers and bundlers but also include chaperones, enzymes, adapters, and receptors whose cellular roles transcend that of structural proteins (Fig. 1). In fact, IF networks now might be considered as signaling scaffolds upon which associated proteins are organized and regulated to control metabolic activities and maintain cell homeostasis. Interactions between nuclear lamins and nuclear-associated proteins are not only required for the normal organization of the karyoskeleton, but have been implicated in regulating gene transcription and cell proliferation. This latter topic is sufficiently complex that we will not cover it in detail in this Chapter, but will refer the reader to other excellent recent reviews in this area (see Goldman *et al.*, 2002; Herrmann and Foisner, 2003; Zastrow *et al.*, 2004). Finally, in some cases, there is a fine line between IFs and IFAPs. Certain IFs harbor extended C-terminal tails that project from the surface of the filament, providing connecting arms that associate with neighboring

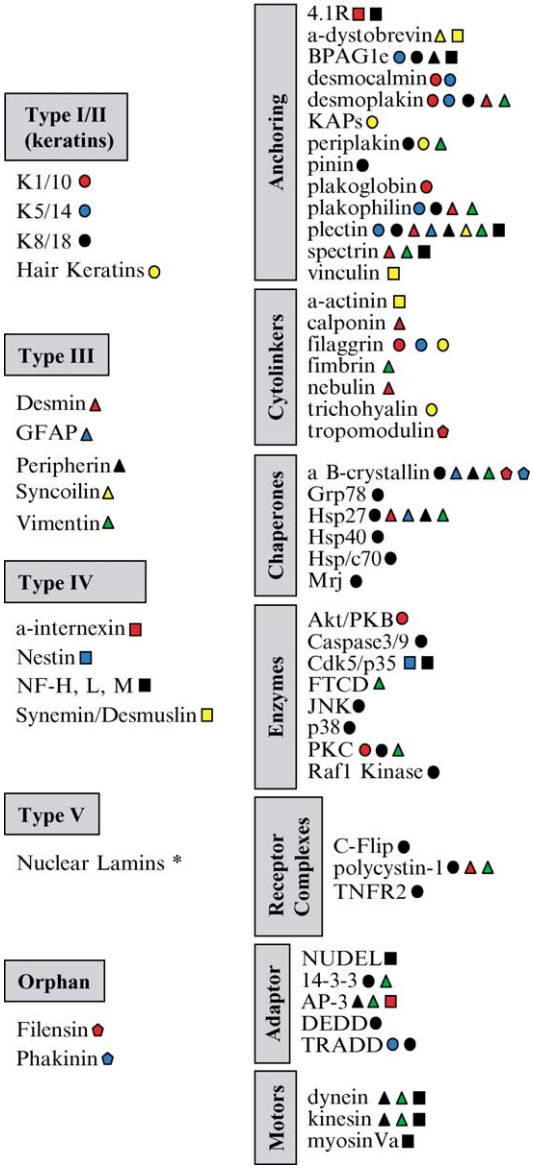


FIG. 1. Diagram of intermediate filaments (IFs) and their associated proteins. Various IFs on the left are color- and shape-coded and matched with interacting partners on the right. For interactions of nuclear lamins with their associated proteins refer to (Zastrow *et al.*, 2004). This figure is not meant to be comprehensive but is rather a summary of those interactions highlighted in this review.

filaments. In some cases, IF polypeptides that coassemble at substoichiometric levels with partner proteins are thought to play critical roles in anchoring filaments to other structures and cellular membranes. In the current Chapter, we will focus on cytoplasmic IFs and their associated proteins; the latter are both organizers of IFs and targets for regulation by this highly diverse class of cytoskeletal polypeptides.

II. Plakin Family

One major family of IFAPs was first recognized in 1991 and subsequently emerged as a gene family providing crosslinking functions for all cytoskeleton networks and the plasma membrane. These proteins are now collectively referred to as the *plakin family* (Jefferson *et al.*, 2004; Leung *et al.*, 2002; Ruhrberg and Watt, 1997), a name which originally derived from the location of many of these proteins in the electron-dense cytoplasmic plaque of cytoskeletal anchoring junctions. It is now known that the extended family includes modular molecules built from building blocks that together comprise enormous proteins. Plakins can contain IF-binding, actin-binding (ABD), coiled-coil rod, spectrin repeat-containing, and growth arrest specific protein 2 (GAS2) MT binding domains (GAR domain), as well as domains that target certain members to junctions such as desmosomes or hemidesmosomes (Jefferson *et al.*, 2004) (Fig. 2). The

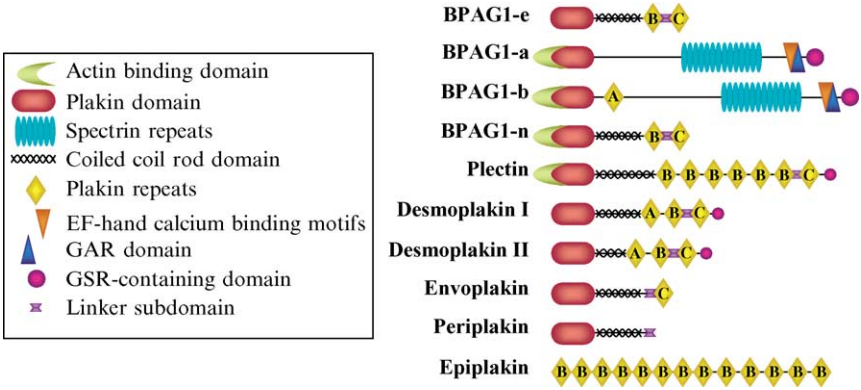


Fig. 2. Domain structure of plakin family members are shown on the right, along with a legend for each of the component domains. A single generic plectin isoform is shown here. Refer to text for a discussion of other plectin isoforms that primarily affect the N-terminus upstream of the ABD.

largest plakins contain spectrin repeats and have now been dubbed *spectraplakins* (Roper *et al.*, 2002).

The most ancient plakin genes, found in the invertebrates *C. elegans* and *Drosophila*, are called VAB-10A/B and SHORTSTOP (kakapo), respectively, and play important roles in adhesion and tissue morphogenesis. These primitive versions of the plakins are highly modular in character, suggesting that several of the more recent epithelial versions were streamlined for specialized functions, primarily as IF-binding cytoskeletal crosslinkers. Interestingly, *Drosophila* does not even have cytoplasmic IFs; the regions most closely resembling IF-binding domains have other functions in junction localization (Roper and Brown, 2003), suggesting either that IF-binding functions were lost or that these domains were later tailored for IF binding. In this review, we will focus on those plakin family members that interact with IFs, beginning with the founding members of the family: desmoplakin, BPAG1, and plectin. We will then move on to discuss the more recently identified envoplakin and periplakin molecules, and the unusual epiplakin protein.

The majority of plakins have been localized to anchoring junctions that mediate either intercellular or cell-substrate adhesion, known as desmosomes and hemidesmosomes, respectively (Fig. 3) (for reviews refer to Borradori and Sonnenberg, 1999; Gaudry *et al.*, 2001; Getsios *et al.*, 2004; Godsel *et al.*, 2004; Jones *et al.*, 1998; Koster et al., 2004a). These junctions share a basic blueprint including transmembrane components that interact either with extracellular matrix molecules (hemidesmosomes) or transmembrane partners in adjacent cells (desmosomes). In hemidesmosomes, the transmembrane components are members of the integrin family as well as BPAG2/BP180; in desmosomes, the desmosomal cadherins (desmogleins and desmocollins) serve this role. In each type of junction, the cytoplasmic tails of these transmembrane molecules act as a scaffolding for the organization of a protein complex that anchors IFs to the plasma membrane. Specialized cell-cell and cell-substrate junctions in endothelial cells, which do not assemble desmosomes or hemidesmosomes, have also recently emerged as important contact sites for IFs (Tsuruta and Jones, 2003; Vincent *et al.*, 2004) (Fig. 3B, D). In each case, plakin family members serve as the IFAP component of the complex. Their critical role is underscored by the existence of human inherited diseases and engineered animal molecules that inactivate their tethering function, leading to tissue fragility diseases (Godsel *et al.*, 2004; Jonkman, 1999; Koster *et al.*, 2004a; McMillan and Shimizu, 2001). In the following section, each of these anchoring junction-associated cytolinkers will be discussed. The reader is referred to recent reviews for details on other plakin family members (Jefferson *et al.*, 2004).

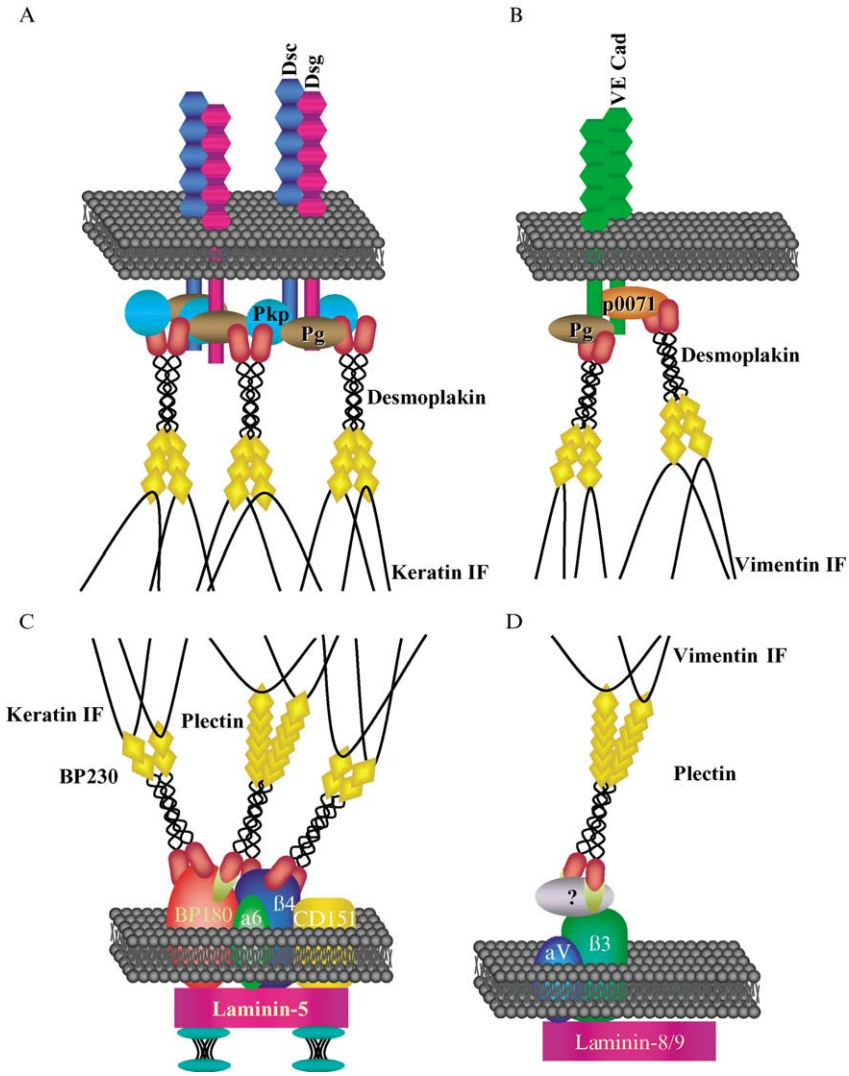


FIG. 3. A comparison of cell-cell and cell-substrated anchoring junctions. (A) Desmosomes anchor keratin filaments through desmoplakin. One half of a desmosome, which is an intercellular junction that anchors intermediate filaments (IFs) to the plasma membrane, is shown. The transmembrane desmosomal cadherins, desmogleins (Dsg) and desmocollins (Dsc), mediate adhesion through their extracellular domains, and associate with plakophilins (Pkp) and plakoglobin (Pg) through their cytoplasmic domains. These proteins in turn interact with the N-terminus of the plakin family member desmoplakin, which anchors IF to the junction through its C-terminus. (B) Endothelial VE-cadherin-based junctions anchor vimentin through

A. *Desmoplakin*

Desmoplakins (DP) I and II were first identified biochemically as abundant components of enriched fractions of desmosomes (Skerrow and Matoltsy, 1974) and were subsequently found to be ubiquitous constituents of these organelles present in all epithelial tissues, heart, mengingeal cells, and follicular dendritic cells of lymph nodes. DP is also found outside of desmosomes in specialized endothelial junctions (Schmelz and Franke, 1993; Schwarz *et al.*, 1990; Vincent *et al.*, 2004). DP I and II are encoded by alternatively spliced products derived from a single gene that gives rise to 332 and 260 kDa proteins, respectively. DPII is missing a 599-amino acid region of the rod domain, but contains N- and C-termini identical to DPI, and appears to have a more limited tissue distribution (Angst *et al.*, 1990). Predictions from sequence analysis as well as empirical data suggest that DP is a dumbbell-shaped molecule with a central α-helical coiled-coiled rod domain, flanked by globular N- and C-terminal domains (Fig. 2) (Green et al., 1990, 1992a,b; O'Keefe *et al.*, 1989; Virata *et al.*, 1992). This basic structure is found in several other plakins, including the 230 kDa hemidesmosomal protein bullous pemphigoid antigen (BPAG1) and the cytoskeletal linking protein, plectin (Green *et al.*, 1992b; Sawamura *et al.*, 1991; Wiche *et al.*, 1991), as well as the cell envelope and junction proteins envoplakin and periplakin (Ruhrberg and Watt, 1997; Ruhrberg *et al.*, 1996, 1997) (Fig. 2). A more recently described family member, epiplakin, contains an extended region homologous to the C-terminus of other plakins but lacks the N-terminal and rod domains.

Analysis of DP cDNA sequence demonstrated that the DP C-terminus comprises three subdomains each made up of 4.5 copies of tandemly

---

desmoplakin. Instead of desmosomes, endothelial cells (Vincent *et al.*, 2004) have specialized intercellular junctions that anchor both actin and vimentin IFs to VE-cadherin-based plasma membrane domains. Vimentin interactions are mediated by desmoplakin, which associates with the VE-cadherin tail via the armadillo proteins plakoglobin and p0071. (C) Hemidesmosomes anchor keratin filaments through plectin and BPAG1. In epithelial cells, hemidesmosomes anchor keratin IFs to the cell substratum. Transmembrane $\alpha 6\beta 4$ integrin (the receptor for laminin 5) and BPAG2/BP180 associate with the plakin family members plectin and BPAG1, which both tether keratin to the plaque through their C-terminal IF-binding domains. The tetraspanin CD151 is thought to play a key role in the assembly of these structures (Nievers *et al.*, 1999). (D) Vimentin-anchoring contacts in endothelial cells. Instead of hemidesmosomes, endothelial cells have specialized focal contact-like structures that anchor vimentin to $\alpha v\beta 3$ integrin complexes at the cell substratum. Although the precise organization of these structures has not been elucidated, it is likely that the plakin family member plectin anchors IFs to these junctions, which also contain proteins such as vinculin that are found in actin-based focal contacts.

organized 38-residue repeats with intervening or linking segments (Green *et al.*, 1990). It was hypothesized, based on the periodicity of charged residues in these domains, that the DP C-terminus interacts with IFs. This prediction was borne out by cell culture, *in vitro* overlay, and yeast two hybrid (Y2H) approaches (Bornslaeger *et al.*, 1996; Kouklis *et al.*, 1994; Meng *et al.*, 1997; Stappenbeck and Green, 1992; Stappenbeck *et al.*, 1993, 1994). More recently, X-ray crystallographic analysis of individual subdomains confirmed that each DP repeat (referred to as a plakin or plectin repeat, or PR) has a substructure 38 residues in length. Each PR comprises an 11-residue beta-hairpin followed by two antiparallel α-helices, representing a novel structural fold. Rather than forming an extended structure typical of other well-known repeating structures, the individual repeats fold to form a discrete globular structure called the plakin repeat domain (PRD). The structures of the different PRDs are similar, and the authors speculate that the PRDs in plakin family members are independent units organized like beads on a string.

The C-terminal region containing these PRDs has been shown to associate with IFs from the range of tissues in which desmosomes are found, including simple and complex epithelial keratins, vimentin, and desmin. The precise mechanism by which these interactions occur is still a matter of debate, however. Based on structural studies cited above, it was proposed that conserved charged residues on the surface of the PRD could be involved in functional interactions with IFs. Indeed, each of the PRD domains was shown by cosedimentation analysis to exhibit low affinity binding to vimentin (Choi *et al.*, 2002), and addition of the highly conserved linker did not detectably increase this association. However, another study demonstrated that the intervening highly conserved linker downstream of the ''B'' PRD is critical for binding to vimentin and the simple epithelial keratins K8/18. The latter finding is in line with the observed importance of this region in other plakin members for binding IFs (Nikolic *et al.*, 1996). Unfortunately, the linker region was not part of the crystal structure and future studies will be required to sort out the relative importance of these different sequences.

Studies have also pointed out that there is sequence specificity in DP interactions with various IFs (Fontao *et al.*, 2003; Meng *et al.*, 1997; Stappenbeck *et al.*, 1993). Although the details of these studies differ, collectively they highlight the importance of the ''C'' PRD plus the last 68 residues for DP's association with keratins and the ''B'' PRD plus linker for type III IF interactions. In addition, phosphorylation of serine 2849, which is 23 residues from the DP C-terminus, impairs interactions with IFs and may be involved in regulating the size of the DP pool that is

competent to incorporate into assembling desmosomes (Fontao *et al.*, 2003; Meng *et al.*, 1997; Stappenbeck *et al.*, 1994).

The identity of IF sequences required for mediating association with the DP C-terminus is also a matter of some debate. Earlier *in vitro* overlay and yeast two hybrid analysis suggested that the presence of a single type II epidermal keratin, either K5 (Kouklis *et al.*, 1994) or K1 (Meng *et al.*, 1997), is sufficient for mediating an association with DP. A region comprising the KSIS motif present in the nonalphahelical head domain of the type II epidermal keratin K5 was implicated as critical for this interaction (Kouklis *et al.*, 1994). The later studies suggested that in the case of K8/18, both partners must be present for a productive interaction with DP, and implicate the importance of the rod in this association (Meng *et al.*, 1997). However, more recent Y3H analysis suggest that the presence of the keratin heterodimer is critical, not just for K8/18 but also for K5/K14, thus pinpointing the rod domain as the most likely partner for DP (Fontao *et al.*, 2003). Together, these results support the idea that the DP C-terminus utilizes distinct sets of overlapping sequences to tailor binding to cell type specific IFs.

The DP N-terminus is required for incorporation into the desmosomal plaque (Bornslaeger *et al.*, 1996; Smith and Fuchs, 1998; Stappenbeck *et al.*, 1993). It contains a region that is conserved throughout all plakin family members, with the exception of epiplakin, now referred to as the *plakin domain*. This domain was originally predicted to form a globular structure, characterized by heptad repeats forming a series of antiparallel bundles. More recently, this prediction has been refined to describe a series of spectrin repeats surrounding a putative SH3 domain (Jefferson *et al.*, 2004). Coimmunoprecipitation, *in vitro* binding, and yeast two hybrid studies have demonstrated that the plakin domain contains binding sites for the armadillo proteins plakoglobin, plakophilins 1-3, and p0071, and may also bind directly to the desmosomal cadherins desmoglein 1 and desmocollin 1. The multiplicity of possible interactions between the DP N-terminus and other junction components provides a mechanism for finetuning the architecture—and accordingly the function—of the junctional plaque (Bonne *et al.*, 2003; Bornslaeger *et al.*, 2001; Calkins *et al.*, 2003; Chen *et al.*, 2002; Hatzfeld *et al.*, 2000, 2003; Kowalczyk *et al.*, 1997, 1999; Smith and Fuchs, 1998; Troyanovsky *et al.*, 1993). High resolution immunogold localization analysis, using antibodies directed specifically against the N- and C-termini of DP, is consistent with their functions in membrane targeting and IF anchoring, respectively, showing that the N-terminus is located closer to the plasma membrane and the C-terminus is deeper in the cytoplasm where IFs are anchored to the plaque (North *et al.*, 1999).

The critical importance of DP throughout development and in adult epidermis and heart is underscored by the existence of both naturally occurring mutations in humans and engineered mutations in mice (for review see Godsel *et al.*, 2004). Complete loss of DP results in an embryonic lethal phenotype at day 6.5 and characterization of these embryos revealed defects in desmosome assembly and association of K8/18 in the embryonic trophectoderm (Gallicano *et al.*, 1998). Animals in which extraembryonic tissues were rescued did survive for several days longer, but eventually succumbed due to defects in heart, neuroepithelium, epidermis, and microvasculature, suggesting the widespread importance of DP in maintaining tissue integrity (Gallicano *et al.*, 2001). Conditional knockout of DP in mouse skin, along with a variety of mutations in humans, results in skin fragility defects sometimes accompanied by cardiomyopathy (Alcalai *et al.*, 2003; Armstrong *et al.*, 1999; Norgett *et al.*, 2000; Rampazzo *et al.*, 2002; Vasioukhin *et al.*, 2001). Although the underlying cellular basis for these tissue defects is still under active study, it has been shown that an engineered mutation uncoupling IFs from desmosomes dramatically decreases the mechanical strength of keratinocyte cell sheets *in vitro*. Less severe, naturally occurring truncations of the DP C-terminus also lead to decreases in intercellular adhesive strength (Huen *et al.*, 2002).

B. BP230 (BPAG1) and its Isoforms

BPAG1 was first identified as a 230 kDa autoantigen in bullous pemphigoid (Stanley *et al.*, 1981), an autoimmune disease characterized by the appearance of large skin blisters. It was subsequently described as a component of hemidesmosomes (Westgate *et al.*, 1985). Together with plectin, BP230 mediates keratin filament association to these cell-substrate anchors (for review see Jones *et al.*, 1998; Koster *et al.*, 2004a; Nievers *et al.*, 1999). Like DP, BPAG1 performs its tethering function by binding to membrane receptors through its N-terminus—in this case, the laminin-5 receptor $\alpha 6\beta 4$ integrin and the transmembrane glycoprotein BP180 (Hopkinson and Jones, 2000; Koster *et al.*, 2003)—and to IF through its C-terminus (Fontao *et al.*, 2003; Leung *et al.*, 2001b; Yang *et al.*, 1996) (Fig. 3C).

Due to its prominent expression in hemidesmosomes, it was predicted that knockout of BPAG1 would result in skin blistering defects in the skin of mice. Indeed, focal blistering in response to stress was observed in mice lacking BPAG1/BP230 (Guo *et al.*, 1995). Surprisingly, however, loss of the BPAG1 locus either by engineered (Guo *et al.*, 1995) or spontaneously occurring mutations in the dystonia musculorum (dt) mouse (Brown

et al., 1995) also led to degeneration of sensory neurons and progressive dystonia in mice.

This provocative phenotype led to the discovery that the BPAG1 locus was larger than originally thought (Sawamura *et al.*, 1990). It encodes additional modules characteristic of the larger plakin family members, giving rise to four distinct isoforms (Fig. 2). The neuronal defects observed in the mutant mice were recently ascribed to loss of BPAG1-a (Leung *et al.*, 2001b), a giant isoform containing calponin homology (CH) domains 1 and 2 (comprising an actin binding domain) whose expression in skeletal muscle may also explain muscle abnormalities observed in dt mice (Dalpe *et al.*, 1999). Somewhat surprisingly, until very recently no corresponding human mutations in BPAG1 had been identified. In a recent report, however, patients with a possible BPAG1 haploinsufficiency with selective disruption of muscle and brain-specific isoforms have been characterized with esophageal atresia and psychomotor retardation (Giorda *et al.*, 2004).

Four known isoforms expressed from the BPAG1 locus result from differential splicing (Leung *et al.*, 2001a; Okumura *et al.*, 2002; Yang *et al.*, 1999). The common structural component of all isoforms is the plakin domain, which mediates binding of BP230 (also known as BPAG1-e) to BP180 (collagen XVII or BPAG2) (Hopkinson and Jones, 2000) and β4 integrin (Hopkinson and Jones, 2000; Koster *et al.*, 2003). The plakin domain in both BPAG1-e and BPAG1-n is followed by a coiled-coil rod domain and two plakin repeat domains B and C, which are connected by a linker subdomain. Like the giant isoforms BPAG1-a (619 kDa) and BPAG1-b (824 kDa), BPAG1n comprises an actin binding domain consisting of two calponin-homology (CH) domains in its outermost N-terminus. Other features of the a and b isoforms are a spectrin repeat-containing rod domain and a microtubule binding domain at the very C-terminus. BPAG1-b, which is expressed in heart and skeletal muscle, also contains a plakin repeat domain, and could theoretically serve as integrator of all three filament types.

The plakin repeat domains of BPAG1 were shown to bind to and colocalize with IF; however, conflicting reports exist about its specificity for different IF types. In one case, data suggested that BPAG1-n colocalizes with neurofilaments, and the BPAG1 plakin repeat domain immunoprecipitated with NF-protein (Yang *et al.*, 1996). Another group reported that the type III neuronal IF protein peripherin interacts with the BPAG1 C-terminus (Leung *et al.*, 1999). The potential ability of the PR containing domain in the ''b'' form to bind IFs awaits confirmation (Jefferson *et al.*, 2004). Recently, yeast three hybrid experiments showed an interaction between the C-terminal domain and the keratin 5/14 heterodimer, but

not keratin 8/18 or vimentin (Fontao *et al.*, 2003), in contrast with the observed broader spectrum of DP binding partners. This difference in the repertoire of BPAG1 and DP partners most likely reflects the wider array of IFs that DP associates with *in vivo* in the desmosomes of heart, meninges, and simple and complex epithelia.

C. Plectin

Plectin and IFAP300 were identified independently as components of high salt/Triton X-100 insoluble, IF-rich preparations derived from cultured cells and tissues (Lieska *et al.*, 1985; Pytela and Wiche, 1980; Wiche *et al.*, 1982; Yang *et al.*, 1985). Along with hemidesmosomal protein 1 (HD1), they have all recently been shown to be identical proteins (Clubb *et al.*, 2000; Herrmann and Wiche, 1987; Okumura *et al.*, 1999). Plectin is present in a wide range of tissues and cells (Wiche, 1989; Wiche *et al.*, 1984, 1983). This versatile plakin functions as a connector between all of the nonmuscle cytoskeletal networks, and as an anchoring protein at plasma membrane specializations including hemidesmosomes, desmosomes, intercalated discs, focal contacts, costameres, and Z-lines of striated muscle sarcomeres (Wiche *et al.*, 1983). In complex epithelia, plectin is concentrated in basal layer cells, reflecting its localization in hemidesmosomes; its function has been clearly demonstrated by engineered and naturally occurring mutations that affect hemidesmosome integrity. But plectin has also been shown to be a component of desmosomes (Skalli *et al.*, 1994), where its presence may be due in part to demonstrated interactions with its fellow plakin member, desmoplakin (Eger *et al.*, 1997). The functional significance of plectin in desmosomes is not as well understood, and patients with plectin mutations do not exhibit obvious desmosome defects.

The cDNAs of human (Liu *et al.*, 1996; McLean *et al.*, 1996) and rat (Elliott *et al.*, 1997; Wiche *et al.*, 1991) plectin each encode an approximately 4600 aa long polypeptide with a predicted molecular mass of over 500 kDa. Structural studies indicate that (like other epithelial plakin family members) plectin consists of two terminal globular domains and a central α-helical coiled-coil rod, and appears as a dumbbell in electron microscopic preparations (Foisner and Wiche, 1987; Wiche, 1989, 1998). Its central alphahelical coiled-coil rod is flanked by an N-terminus comprising CH1, two ABD modules, and a plakin domain, and a C-terminus consisting of six PRD domains containing the conserved linker downstream of the fifth PRD. Like DP, plectin contains a terminal GSR domain that is highly conserved between these two plakin family members. This

region in plectin, but not DP, binds to MT (Sun *et al.*, 2001). The fact that plectin has the ability to associate with actin, MT, and IFs has led to the idea that it functionally integrates all three structural components of the cytoskeleton in cells, a concept dramatically illustrated by electron microscopical analysis of plectin ''sidearms'' physically bridging these elements (Foisner *et al.*, 1995; Svitkina *et al.*, 1996; Wiche, 1998).

A number of studies have dissected the role of specific sequences in conferring cytoskeletal linking functions to plectin. Plectin's C-terminal domain is built from six PRDs and includes the conserved linker region downstream of the fifth repeat. This linker comprises binding sites for vimentin and K8/18, as shown by *in vitro* binding assays and colocalization of wild-type and mutant protein fragments in transfected cells (Nikolic *et al.*, 1996). Plectin also binds to GFAP, all three neurofilament proteins (Foisner *et al.*, 1988), and lamin B (Foisner *et al.*, 1991). Like DP, phosphorylation of plectin regulates association with various IFs. Phosphorylation by protein kinases A and C leads to decreased lamin B binding, whereas interaction with vimentin was enhanced by PKA phosphorylation of plectin (Foisner *et al.*, 1991).

The N-terminal plectin ABD consists of two calponin-homology domains, each comprising four primary α-helices connected by loops and shorter helices. This functional actin-binding site exhibits the same affinity for actin as that of other similar actin-binding proteins (Andrä *et al.*, 1998; Fontao *et al.*, 2001; Geerts *et al.*, 1999). The N-terminal actin binding and plakin domains are also involved in binding to β4 integrin, an association that competes for actin binding (Geerts *et al.*, 1999; Koster *et al.*, 2004b). An additional binding site for β4 integrin is found in the final PRD (Rezniczek *et al.*, 1998). X-ray crystallographic analysis of the plectin ABD revealed that it is folded similarly to other ABDs (Garcia-Alvarez *et al.*, 2003; Sevcik *et al.*, 2004). The structure formed by two calponin homology domains exhibits a closed structure that undergoes a conformational change upon binding of actin, but not β4 integrin. This latter observation may help explain how plectin competitively binds F-actin and $\alpha6\beta4$ (Garcia-Alvarez *et al.*, 2003).

The study by Sevcik *et al.* (2004) also reported an additional vimentin binding site in the first CH-domain of plectin. This region shows high homology to fimbrin, another CH domain containing protein that associates with IFs, as described in more detail below. Unlike the C-terminal IF binding site, this N-terminal site does not bind to filamentous vimentin as judged by cosedimentation experiments (Sevcik *et al.*, 2004). Based on these data, the authors propose that this additional N-terminal IF binding site of plectin may regulate dynamics, rather than stabilization of the IF network.

The plectin gene contains more than 40 exons encompassing over 32 kb on chromosome 15, generating an unusual 5′ transcript complexity and giving rise to many plectin isoforms (Rezniczek *et al.*, 2003). Sixteen alternatively spliced exons have been identified, 11 of them directly spliced into exon 2 encoding the ABD (Fuchs *et al.*, 1999; Rezniczek *et al.*, 2003). However, in several cases, differential splicing results in retention of only one CH domain. An isoform lacking exon 31, which encodes most of the rod domain, was found for rat (Elliott *et al.*, 1997) and is presumed for human plectin (Schröder *et al.*, 2000). These isoforms show tissue, cell type, and developmental stage variation in expression that presumably reflects their multiple and diverse functions (Elliott *et al.*, 1997; Fuchs *et al.*, 1999; Rezniczek *et al.*, 2003). On the cellular level, some of the isoforms align with actin stress fibers when expressed in cultured cells (Rezniczek *et al.*, 2003); others show divergent patterns. For example, the recombinant plectin isoform 1f was shown to localize to focal contacts (Rezniczek *et al.*, 2003). Thus, it could serve as a potential anchor of IFs to the plasma membrane at the site of focal contacts in fibroblasts and, possibly, endothelial cells (Gonzales *et al.*, 2001; Seifert *et al.*, 1992) (Fig. 3D).

As described above, both the N-terminal plakin and C-terminal domains of plectin bind to the cytoplasmic tail of the $\beta 4$ integrin component of the hemidesmosome (Rezniczek *et al.*, 1998). The importance of plectin to the structure and function of the hemidesmosome is indicated by the consequences of its loss in both humans and mice. Although plectin is dispensable for the formation of hemidesmosomes per se, in skin of plectin (−/−) mice ablation of plectin results in skin blistering due to hemidesmosome perturbation (Andrä *et al.*, 1997; Pulkkinen *et al.*, 1996; Uitto *et al.*, 1996). The plectin isoform 1a was shown to localize to hemidesmosomes in rat skin preparations (Rezniczek *et al.*, 1998) and to rescue hemidesmosome-like anchoring contacts in plectin-deficient keratinocytes (Andrä *et al.*, 2003). Moreover, mice and humans lacking plectin also exhibit muscle deficiencies. This is not surprising, since plectin is a component of Z-disks of striated muscle, dense plaques of smooth muscle cells, and intercalated disks of cardiac muscle (Wiche *et al.*, 1983). Indeed, in striated muscle, plectin serves not only as a linker of desmin to the Z-disks (Hijikata *et al.*, 1999), but also plays a role in extending the desmin IF network to the costameric dense plaques (Hijikata *et al.*, 2003) where it may bind to vinculin through actin. Interestingly, Fuchs *et al.* (1999) found that the isoform prevailing in muscle, which contains the differentially spliced exon 2α, has greater actin binding activity than the isoform lacking this small exon.

Most recently, it has been proposed that plectin may function as a signaling scaffold. A yeast two hybrid screen uncovered the PKC anchoring protein, RACK1, as a novel interacting protein. The authors went on to show that EGF stimulates the formation of a plectin-RACK1-PKC complex, coinciding with a redistribution of RACK from the cytoskeleton fraction to the soluble pool (Osmanagic-Myers and Wiche, 2004). In plectin null cells, RACK1 was mislocalized, exhibiting a reduction in cytoskeletal association. The authors concluded that plectin may function to sequester RACK on the cytoskeleton in uninduced cells and then to shuttle it to other locations, possibly to membrane sites, upon EGF induction.

D. Envoplakin and Periplakin

Envoplakin and periplakin were first identified as 210 and 195 kDa transglutaminase substrates that were expressed at higher levels during keratinocyte differentiation, and became crosslinked into the assembling cornified cell envelope (Simon and Green, 1984). Both envoplakin and periplakin were originally reported as present in complex epithelial tissues but not most simple epithelia, mesenchymal tissues, or the heart. More recent studies suggest that periplakin is expressed more widely in a number of simple epithelia as well (Kazerounian *et al.*, 2002).

Like other epithelial members of the plakin family, envoplakin and periplakin are built from a plakin-domain containing N-terminus, a central alphahelical coiled-coiled rod domain, and a C-terminus (Ruhrberg and Watt, 1997; Ruhrberg *et al.*, 1996, 1997) (Fig. 2). In contrast to other plakin family members, both envoplakin and periplakin have relatively short C-termini. Envoplakin contains the so-called linker domain, followed by a single PRD and short tail, whereas periplakin contains only the ''linker.'' Also in contrast to other plakin family members, the distribution of acidic and basic residues in the rod domains of envoplakin and periplakin suggest that they may not only be capable of homodimerizing, but also heterodimerizing. This prediction was supported by coimmunoprecipitation studies showing that the proteins exist in a complex (DiColandrea *et al.*, 2000). Recent biochemical analysis revealed that envoplakin and periplakin form soluble complexes with equimolar stoichiometry, and that envoplakin requires the presence of periplakin for its solubility. Further chemical crosslinking analysis and electron microscopical analysis showed that the primary soluble form of the envoplakin/periplakin rod domains is a dimer, and that full-length proteins form higher-order oligomers (Kalinin *et al.*, 2004).

Both envoplakin and periplakin have been localized to desmosomes, but their distribution is different from other desmosome markers in that both envoplakin and periplakin radiate outwards underneath the plasma membrane from the plaque proper. Periplakin and envoplakin have been shown to participate in the scaffolding of the cornified envelope, becoming crosslinked by transglutaminases to other cell envelope and desmosome proteins (Steinert and Marekov, 1999) as well as lipids (Marekov and Steinert, 1998). Both periplakin and envoplakin/periplakin oligomers bind synthetic lipid vesicles, whose composition is similar to the cytoplasmic face of eukaryotic cell plasma membranes. These data support the purported role of these plakin family members as scaffolding proteins for cornified cell envelope assembly (Kalinin *et al.*, 2004).

In spite of their shorter C-termini, envoplakin and periplakin appear to associate with IFs when coexpressed in cells (DiColandrea *et al.*, 2000), a property that depends on their linker domains. Subsequent Y2H and overlay studies support a role for the periplakin (but not envoplakin) linker, in direct association with K8 and vimentin (Karashima and Watt, 2002; Kazerounian *et al.*, 2002). It has been proposed that periplakin may stabilize envoplakin's interaction with the IF network (Karashima and Watt, 2002). Periplakin may also be necessary for localizing envoplakin to cell-cell contact sites, and in addition, appears to harbor a domain in its N-terminus that facilitates association with the actin cytoskeleton (DiColandrea *et al.*, 2000). As neither envoplakin nor periplakin harbor a CH-like ABD in their N-termini, the question remains as to how this association is mediated. However, a recent study has identified kazrin as a novel periplakin-interacting protein whose binding site maps to the periplakin N-terminus. Kazrin's distribution overlaps with that of desmosomes, but it is also associated with the membrane cortex in other interdesmosomal regions, and thus seems like a good candidate for participating in the anchorage of periplakin to cortical actin (Groot *et al.*, 2004).

Like other plakin family members, envoplakin and periplakin autoantibodies are also present in patients with the autoimmune disease paraneoplastic pemphigus, although it is not considered that this component is pathogenic in the disease (Kiyokawa *et al.*, 1998; Mahoney *et al.*, 1998). Whereas mice lacking periplakin exhibited no detectable phenotype (Aho *et al.*, 2004), ablation of the envoplakin gene led to a relatively minor phenotype, including a delay in cornified envelope formation and acquisition of barrier function (Maatta *et al.*, 2001). This mild phenotype is in striking contrast to that observed in animals deficient in desmoplakin or plectin; therefore, it seems likely that envoplakin and periplakin work coordinately with other proteins that play compensatory roles in the knock out animals.

E. *Epiplakin*

Epiplakin is an unusual plakin family member encoded by a single remarkably large exon (>20 kb) giving rise to a protein consisting of 16 PRD domains linked by intervening sequences that do not harbor a conserved linker sequence (Spazierer *et al.*, 2003).The protein was originally identified by autoantibodies in patients with an autoimmune skin disease (Fujiwara *et al.*, 1996) and appears to be restricted to epithelial tissues (Fujiwara *et al.*, 2001; Spazierer *et al.*, 2003). Although it has been purported as having a role in bundling IFs, and this is an appealing speculation based on its unusual structure, very little has been done thus far to characterize the function of this newest plakin family member.

III. Cell Surface–IF Interactions

Plakins are not the only proteins that tether IFs to the cell surface. A number of other proteins localized to cell surface structures including desmosomes, focal contacts, and muscle costameres also contribute to IF anchorage at plasma membranes. IFs may also associate with cell surface receptors outside of ultrastructurally distinct structures, a topic that will be dealt with below.

A. *Proteins in Desmosomes*

1. *Armadillo Proteins: Plakophilins and Plakoglobin*

Plakophilins and plakoglobin are both members of the larger armadillo gene family of dual junctional and signaling molecules. This family is named after the *Drosophila* segment polarity gene armadillo, later found to be the invertebrate orthologue of β-catenin (Peifer *et al.*, 1992). This latter protein not only plays a key role as an adapter in adherens junctions but also as a downstream effector in the Wnt/wingless signaling pathway, which governs developmental patterning in the fly and cell fate decisions in vertebrates (Nelson and Nusse, 2004). Its central domain consists of a series of 12 imperfect repeats, termed *armadillo repeats*, that provide scaffolding for interaction with binding partners involved in both junction structure and transcriptional regulation. In addition to β-catenin, members of this family include the junctional molecules plakoglobin, p120 catenin, and plakophilins (Anastasiadis and Reynolds, 2000; Hatzfeld, 1999) (Fig. 3).

Plakoglobin is β-catenin's nearest relative, and is a constituent of both adherens junctions and desmosomes (Zhurinsky *et al.*, 2000). Its central

armadillo domain binds to both classic and desmosomal cadherins, and its shorter flanking N- and C-terminal domains serve regulatory functions. The C-terminus in particular appears to control desmosome size, as its deletion results in formation of laterally extended desmosomes. This region of plakoglobin also contains a number of phosphorylation sites that govern protein interactions (Gaudry *et al.*, 2001; Miravet *et al.*, 2003; Palka and Green, 1997). Plakoglobin is not generally considered to be a major direct binding partner for IFs, but rather an indirect link through interaction of its armadillo repeats with the IFAP desmoplakin. However, one report suggested that plakoglobin binds weakly to epidermal keratins, as shown by *in vitro* overlay assays (Smith and Fuchs, 1998) and thus in certain cases it may contribute more directly to IF binding.

More substantial data supports an interaction between plakophilins (PKPs) and IFs. This subfamily of the armadillo family is most similar to the adherens junction protein p120catenin in its structure and organization of central armadillo repeats. There are three major desmosomal PKPs (1–3) (reviewed in Anastasiadis and Reynolds, 2000; Hatzfeld, 1999). The PKPs exhibit cell type specific differences in their tissue distribution and protein interaction partners, but collectively they associate with most desmosome proteins including plakoglobin, desmoplakin, and the desmosomal cadherin tails (for review see Godsel *et al.*, 2004). Like other armadillo family members, the PKPs localize both to junctions and the nucleus, although their function in the nucleus is not well understood. Unlike plakoglobin, all of the binding sites for desmosome proteins identified so far lie in the extended N-terminal domain, rather than the central armadillo domain (Bonne *et al.*, 2003; Chen *et al.*, 2002; Hatzfeld *et al.*, 2000; Kowalczyk *et al.*, 1999).

PKP1 (formerly known as Band 6 of desmosome-enriched preparations) was first identified biochemically as a keratin-binding protein (Kapprell *et al.*, 1988) and later identified by sequence analysis to be a member of the armadillo family (Hatzfeld *et al.*, 1994). More recent studies using yeast two hybrid and *in vitro* solid assays demonstrated that PKP1a and PKP2a bind to simple and epidermal keratins and, to some extent, vimentin (Hatzfeld *et al.*, 2000; Hofmann *et al.*, 2000). PKP1 in particular binds avidly to IFs assembled *in vitro* from CKs 8/18, 5/14, vimentin, or desmin and facilitates formation of thick IF bundles. PKP3 has also been shown to associate with K18 (Bonne *et al.*, 2003). The IF binding site on PKP maps to the N-terminal head domain, raising the question of how this domain could associate with both the membrane complex and IFs at the same time. Indeed, the physiological relevance of these interactions has been called into question based on the fact that PKP isn't normally seen to associate with IFs in cells, and its position in the desmosome plaque suggests it

may be inaccessible to IFs (North *et al.*, 1999). However, it has been suggested that in DP-deficient or mutant cells, PKP may participate in tethering keratin filaments to the plaque (Norgett *et al.*, 2000; Vasioukhin *et al.*, 2001).

2. *Pinin*

Pinin was identified as a 140 kDa protein widely expressed in a number of epithelial tissues, brain, and heart, which incorporates into mature desmosomes near where keratin filaments impinge upon the desmosomal plaque (Ouyang and Sugrue, 1996). Pinin contains three coiled-coil domains, a small stretch of glycine loops, a short glutamine-proline-rich domain and a polyserine domain (Ouyang *et al.*, 1997). Yeast two hybrid and *in vitro* overlay analysis have demonstrated that the N-terminal domain of pinin interacts with the 2B rod domain of keratins 18, 8, and 19 (Shi and Sugrue, 2000). Interestingly, another 140 kDa protein, identified years ago as a lamin B-like protein present in small amounts in preparations of bovine muzzle desmosomes, was shown to associate with vimentin in an *in vitro* overlay assay (Cartaud *et al.*, 1990). Any relationship to pinin has not been proven, however.

3. *Desmocalmin/Keratocalmin*

Desmocalmin, which may be the bovine equivalent to a human protein called keratocalmin (Fairley *et al.*, 1991), was identified in the mid-1980s as a desmosome protein that binds both to calmodulin in a Ca^{2+}-dependent manner and to reassembled keratin filaments *in vitro* in the presence of Mg^{2+} (Tsukita and Tsukita, 1985). Unfortunately, due to the loss of critical reagents, progress on this potentially interesting regulator of IF binding has been impeded, and further work will be necessary to revisit its potential role in desmosome assembly and regulation.

B. *Polycystin-1*

Polycystin-1 is a widely expressed, large, multispan membrane protein that is a target for mutation in most patients with autosomal dominant polycystic kidney disease. Mutations in this protein interfere with the normal differentiation of renal tubular epithelial cells, resulting in the formation of large cystic kidneys. Although the mechanisms underlying these mutant phenotypes are poorly understood, the polycystin C-terminal domain has been implicated in modulating wnt signaling by stabilizing the armadillo protein β-catenin (Kim *et al.*, 1999b). It may also regulate G-protein signaling by similarly stabilizing a regulator of G-protein

signaling (RGS), RGS7, that is normally rapidly degraded by the proteosome pathway (Kim *et al.*, 1999a).

To gain further insight into the functions of polycystin-1, the C-terminal tail (P1CT) was used to screen a kidney yeast two hybrid library. In this study, vimentin, K8/K18, and desmin were identified as binding partners, and these interactions were confirmed by GST pull down and *in vitro* polymerization assays (Xu *et al.*, 2001). The polycystin–IF interactions were shown to be mediated by coiled-coil regions in both interacting partners. In addition to associating with the IF cytoskeleton, this versatile protein has been shown to specifically colocalize with IF-associated desmosomes (Scheffers *et al.*, 2000; Xu *et al.*, 2001) and to form a complex with E-cadherin/β-catenin in kidney cells. This latter interaction is impaired in cells from patients with polycystic kidney disease, apparently due to increased phosphorylation of polycystin-1 (Roitbak *et al.*, 2004). Polycystin-1 is also localized to focal contacts where it closely associates with $\alpha 2\beta 1$ integrin (Wilson *et al.*, 1999). It is still unclear whether polycystin utilizes cytoskeletal-cell junction scaffolding, properly localizing itself to perform signaling functions that govern kidney cell differentiation and polarization, or whether these associations are more directly involved in disease etiology (or both).

C. Association of IFs with the Actin-Rich Cortex

The cortical region of many cells is enriched in actin and associated actin-binding proteins, which function in motility, cell shape maintenance, and membrane protein distribution in polarized cells. In some cases, discrete structures anchor actin to the membrane, as is the case for intercellular adherens junctions and cell-substrate focal contacts. In certain special cell types, the fundamental blueprint for an adherens junction is taken to a new structural level, serving as scaffolding for cell-type specific complexes, such as the dystrophin-associated protein complex (DPC) in striated muscle. Although for years morphological studies have described a close association with IF with the actin-rich cortex, recent advances in methods to study protein–protein interactions have provided new insight into the intimate structural and functional relationship between IFs and these membrane domains.

1. Focal Contacts

Fibroblasts and endothelial cells do not assemble hemidesmosomes. Rather, their matrix adhesive devices are so-called *focal contacts*. Pivotal components of focal contacts are heterodimeric integrins that bind to a

variety of matrix molecules including fibronectin, vitronectin, and laminins (Brakebusch and Fassler, 2003; Hynes, 2002; Zamir and Geiger, 2001). Classically, integrins associate with cytoplasmic actin microfilaments through a number of scaffold and linker proteins. However, several reports indicate that focal contacts associate also with the vimentin IF network (Bershadsky *et al.*, 1987; Gonzales *et al.*, 2001). Although not definitively identified as a linker for vimentin at focal contacts, plectin is a likely candidate, particularly the murine plectin isoform 1f (Rezniczek *et al.*, 2003). Another potential connector for vimentin in integrin-based contacts is MICAL, a HEF1/CasL interacting protein (Suzuki *et al.*, 2002). HEF1/CasL is a member of the Cas family, whose members are docking proteins involved in integrin signaling with a broad range of cellular processes (O'Neill *et al.*, 2000). Its expression is highest in T cells, B cells, and epithelia (O'Neill *et al.*, 2000), but it is also present in fibroblasts and focal contacts of adherent cells (Law *et al.*, 2000).

Are vimentin-associated matrix adhesive devices in nonepithelial cells hemidesmosome equivalents? Several lines of evidence support this possibility. Although vimentin knockout mice show no obvious defects in development, cells derived from those mice display several defects. Wound healing defects in vimentin knockout mice are thought to be caused by failure of mesenchymal collagen contraction (Eckes *et al.*, 2000). Vimentin-deficient fibroblasts show impaired mechanical stability, motility, and directional migration, leading to the proposal that the lack of vimentin induces aberrant organization of focal contacts in the null cells (Eckes *et al.*, 1998) Experimental support for this proposal comes from studies in cultured endothelial cells in which vimentin knockdown resulted in reduced focal contact size, even though the microfilament and microtubule cytoskeleton systems show no obvious irregularities (Tsuruta and Jones, 2003). Furthermore, when exposed to fluid shear stress, the vimentin-deficient cells showed impaired adhesion to the substrate. Intriguingly, hemidesmosomes lacking a connection to keratin are also reduced in size (Gache *et al.*, 1996; Guo *et al.*, 1995). Collectively, these observations suggest that vimentin IFs may regulate cell adhesion in a manner parallel to that of keratins, which stabilize hemidesmosome adhesion to the basement membrane (Jones *et al.*, 1998).

2. *Muscle Cell Surface Attachments for IFs*

In striated skeletal and cardiac muscle, the desmin IF network forms an exosarcomeric lattice surrounding the myofibrils. Parallel filaments link Z-disks within a myofibril, and a prominent series of transverse filaments link neighboring myofibrils at the level of the Z-disks (Price and

Sanger, 1979). In addition, transverse IFs extend out to the plasma membrane or sarcolemma, where they interact with cortical structures called costameres that act as sites of force transmission (Danowski *et al.*, 1992; Pardo *et al.*, 1983). *Costameres* are cortical densities comprising numerous proteins that circumferentially line up in register with the Zdisks of peripheral myofibrils and link the force-generating sarcomeres with the plasma membrane (Ervasti, 2003) (Fig. 4). The desmin network is critical for maintaining muscle structure and function in response to stress, as loss or mutation leads to cardiomyopathy and muscle dystrophy in highly used skeletal muscles (Carlsson and Thornell, 2001).

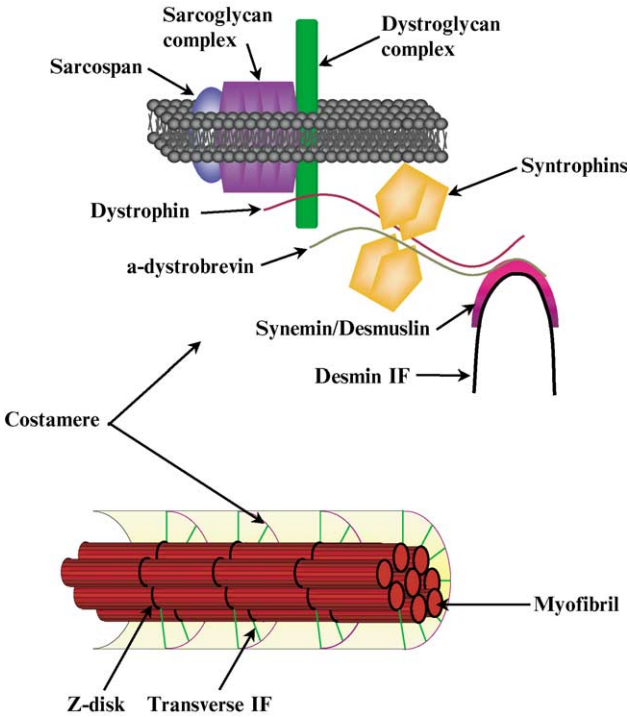


FIG. 4. Structure of striated muscle costameres and the DPC. A single membrane-associated costamere from a portion of a striated muscle fiber is magnified above to show the components of the dystrophin-associated protein complex that are involved in linking desmin intermediate filaments (IFs) to the muscle cell membrane. Additional actin-associated proteins present at these sites (including vinculin, talin, spectrin, and ankyrin) are not shown here. In addition to components of the DPC, plectin has also been localized to costameres, and likely contributes to linking desmin IFs to actin-associated structures.

Costameres have been compared to focal contacts or adherens-type junctions that, in fibroblasts, link actin to the membrane. Like focal contacts, costameres contain actin-associated proteins such as vinculin, talin, spectrin, and ankyrin (Craig and Pardo, 1983; Small *et al.*, 1992). In addition, they act as major integrators of the striated muscle IF cytoskeleton. In smooth muscle, dense plaque material forms riblike arrays over the cell surface (Small, 1995) that, like costameres, provide anchorage sites for actin and desmin IFs (Small and Gimona, 1998). One of the key assemblies localized to the costamere is the dystrophin-associated protein complex (DPC), a collection of structural and signaling proteins critical for mechanical integrity of muscle. This complex is a frequent target of human mutation, resulting in a variety of muscular dystrophic diseases (Blake and Martin-Rendon, 2002). A key coordinating component in the IF-specific portion of the costamere and the DPC is α-dystrobrevin.

3. α-Dystrobrevin and the IF-Associated DPC

Dystrobrevins are a small family of dystrophin-related proteins encoded by two genes, one of which—α-dystrobrevin—encodes at least three proteins all expressed in cardiac and skeletal muscle. α-Dystrobrevin is a key component of the DPC, and its loss results in neuromuscular junction defects and muscular dystrophy. It is a cytoplasmic protein indirectly linked with the transmembrane sarcoglycan components via dystrophin and other components of the DPC (Fig. 4) (Blake and Martin-Rendon, 2002). Recently, two novel type IV IF proteins—syncoilin (Newey *et al.*, 2001) and desmuslin (Mizuno *et al.*, 2001)—were both discovered in a yeast two hybrid screen for α-dystrobrevin binding partners and shown to associate with this key component of the DPC (Blake and Martin-Rendon, 2002).

Syncoilin is highly expressed in skeletal and cardiac muscle and is localized to the neuromuscular junction, sarcolemma, and Z-lines. Likewise, desmuslin is expressed in heart and skeletal muscle and localized at Z-lines. It was shown that syncoilin and desmin interact directly, but do not coassemble into filaments; in fact, evidence suggests that syncoilin does not participate in filament formation at all. It was proposed that syncoilin helps anchor the desmin IF network at the sarcolemma and the neuromuscular junction (Poon *et al.*, 2002). More recent work has analyzed patients with a desmin-related cardiomyopathy in which patients with desmin accumulation also exhibit an upregulation of syncoilin and accumulation of other elements of the DPC. These defects were correlated with a disappearance of both α-dystrobrevin-1 and neuronal nitric oxide

synthase (nNOS) from the membrane, possibly contributing to ischaemic injury. The authors suggest that the α-dystrobrevin-syncoilin-desmin interaction may be a key contributing factor to loss of α-dystrobrevin from the membrane, and subsequent loss of nNOS (Howman *et al.*, 2003).

The desmuslin sequence is similar to that of synemin (Hesse *et al.*, 2001), a desmin-binding protein originally identified as an IFAP (Granger and Lazarides, 1980), but which now—based on sequence homology comparisons—is considered to be a *bona fide* IF (Bellin *et al.*, 1999). The similarity between desmuslin and synemin raises the possibility that desmuslin is the human orthologue of synemin (Hesse *et al.*, 2001). Two synemin isoforms resulting from the alternative splicing of a single gene have been termed α- and β-synemin, and a third isoform was recently identified in humans (Xue *et al.*, 2004). β-Synemin appears to be the predominant form in striated muscle (Titeux *et al.*, 2001), where it colocalizes with costameres and muscle Z-lines, and is enriched at the neuromuscular and myotendinous junctions (Mizuno *et al.*, 2004). α- and β-synemin levels are comparable in smooth muscle. Synemin has also been shown to interact with myofibrillar Z-line α-actinin and the costameric protein vinculin (Bellin *et al.*, 1999). Thus, desmuslin/synemin may help to link the muscle cell IF desmin to myofibrillar Z-lines and costameres at the muscle plasma membrane (Bellin *et al.*, 1999, 2001).

Naturally occurring human mutations in the DPC are beginning to reveal discrete functions for the component parts of this IF-associated protein meshwork. For instance, whereas disruption of dystrophin leads to muscular dystrophy accompanied by sarcolemmal damage, ablation of α-dystrobrevin results in muscular dystrophy accompanied by minimal membrane damage (Grady *et al.*, 1999). Although the underlying basis of these differences is not fully understood, it is proposed that the α-dystrobrevin-IF network is involved in lateral force transmission through the costamere, while other components are sufficient to maintain membrane integrity (Ervasti, 2003).

Although much of the focus has been on the DPC of striated muscle, it is likely that desmin attachments to dense plaques of smooth muscle play critical roles in regulating the transmission of contractile forces in this tissue as well. This is particularly relevant in light of the observed defects in smooth muscle of desmin-deficient mice, in which active force per cross-sectional area was reduced to 40% of controls of smooth muscle tissue (Sjuve *et al.*, 1998). IFAP candidates for serving this linking function are plectin and other components of the actin-rich cortex, including calponin (which also plays a role in the cytoplasm of smooth muscle cell dense bodies; see below), and the spectrin/ankyrin complex.

4. *Spectrin and Ankyrin*

In addition to ABD-containing proteins such as plectin, ankyrin and spectrin/fodrin (nonerythrocyte spectrin) may contribute to linking IFs to the actin-rich cortex in both muscle and nonmuscle cells. A number of early studies pointed to these proteins as potential IF linkers; however, a caveat of these studies is that many were performed prior to widespread use of recombinant proteins and involved the use of purified or enriched preparations that could have been contaminated with other intermediary molecules. The nonerythrocyte form of spectrin was implicated early on to play such a role when Mangeat and Burridge (1984) demonstrated that microinjection of an antibody directed against spectrin resulted in an altered distribution of vimentin-containing IF in living fibroblasts. However, it remained possible that these reported IF-spectrin interactions could be indirect, as Wiche and colleagues demonstrated that the IFAP plectin interacts with spectrin and the 240 kDa chain of fodrin (nonerythrocyte spectrin), and thus could serve as the link between IF networks and spectrin (Herrmann and Wiche, 1987). A more direct role for spectrin binding to desmin IF was implicated from *in vitro* experiments by Langley and Cohen (1986). These investigators reported that desmin filaments and erythrocyte spectrin cosediment. In addition, electron microscopy was used to demonstrate the interaction of desmin IF via multiple lateral associations with IOV (inside-out membrane vesicles from erythrocytes). However, the possibility of a contaminant such as ankyrin (see below) that might mediate an interaction between spectrin and desmin IF was not ruled out. The same investigators also published *in vitro* data to suggest that spectrins from erythrocytes versus spectrins from the brain bind with differing affinities to vimentin IFs and neurofilaments (Langley and Cohen, 1987). Later experiments went on to demonstrate an interaction between the β-subunit of brain spectrin and the NF-L rod; these investigators made efforts to verify that their spectrin preparations were not contaminated with other minor components, such as ankyrin (Frappier *et al.*, 1991, 1992).

In apparent contrast to the above results, Georgatos and Marchesi (1985) demonstrated that unpolymerized lens vimentin was capable of binding in a saturable manner to human erythrocyte IOVs stripped of actin and spectrin. Removal of ankyrin, a protein that links the actin-spectrin complex to the erythrocyte membrane, significantly lowered binding and a polyclonal antiankyrin antibody blocked 90% of the binding. These investigators went on to demonstrate that vimentin IF subunits associate with human erythrocyte plasma membranes (Georgatos *et al.*, 1985), and desmin and vimentin IF subunits interact with avian

erythrocyte plasma membranes (Georgatos and Blobel, 1987b; Georgatos *et al.*, 1987) via their N-terminal head domains, an interaction that may also involve synemin (Granger and Lazarides, 1982). Affinity chromatography was used to demonstrate an association between vimentin and nuclear lamins, and interestingly, the C-terminal tail of vimentin was shown to be critical for this association (Georgatos and Blobel, 1987a).

More recent work suggested that one of four 4.1 proteins—4.1R—may associate with neurofilament proteins in forebrain postsynaptic densities, thus regulating the associated spectrin-rich cortex (Scott *et al.*, 2001). Blot overlay analyses demonstrated that, in addition to spectrin and actin, postsynaptic density polypeptides included NF-L and α-internexin as interacting partners for 4.1R. Collectively, these studies emphasize that common themes are used in different cell types to both strengthen plasma membrane domains enriched in actin and IF polypeptides and to coordinate these sites with cytoplasmic architecture.

IV. Other Crosslinkers not in the Plakin Family

A. *Keratinocyte Crosslinkers and Bundlers*

1. *Filaggrin*

In contrast to the emerging family of cytolinker proteins that link IFs to other cytoskeletal elements and structures, proteins that bundle IFs are less diverse and seem to be limited to keratin-associated proteins, such as filaggrin and trichohyalin (reviewed in Coulombe *et al.*, 2000). Filaggrin was identified in the early 1980s as highly basic, histidine-rich protein abundant in the upper differentiated layers of the skin (reviewed in Dale *et al.*, 1993; Presland and Dale, 2000; Steinert *et al.*, 1981). It is synthesized in the granular layer as a larger, highly phosphorylated precursor that is stored in keratohyalin granules in the suprabasal, granular layer of keratinizing epithelia (including epidermis, gingiva, and tongue). This precursor form, called *profillaggrin*, consists of tandemly arrayed (in humans, 10–12; Gan *et al.*, 1990) repeating units of the filaggrin monomer, surrounded by partial repeats and nonrepeat sequences at either end. The N-terminus consists of an A-domain comprising two functional calcium-binding EF hands and a cationic B-domain (Markova *et al.*, 1993; Presland *et al.*, 1992, 1995). It has been proposed that this domain may have multiple functions in sequestering calcium (possibly facilitating the calcium-dependent processing of profilaggrin to filaggrin) and in regulating

the solubility of the protein and consequently keratohyalin granule formation.

During terminal differentiation of keratinocytes of skin into stratum corneum, profilaggrin is dephosphorylated and proteolytically processed into single filaggrin proteins of 37 kDa (in humans) by means of several proteases (Pearton *et al.*, 2001; Resing *et al.*, 1995; Yamazaki *et al.*, 1997). Filaggrin then crosslinks keratin filaments into macrofibrils orientated parallel to the surface of the skin (Dale *et al.*, 1978; Manabe *et al.*, 1991; Steinert *et al.*, 1981). Filaggrin binds to the α-helical rod domain of keratin (Dale and Presland, 1999). This interaction is believed to depend solely on ionic interactions between the conserved positive and negative charges on the β-turns of filaggrins on one hand and along the rod domains of IFs on the other (Mack *et al.*, 1993). Filaggrin association with keratin may facilitate disulfide bond formation between keratin polypeptide chains during the remodeling that occurs in upper layers, which in turn may allow their survival during terminal differentiation (Dale *et al.*, 1994; Presland and Dale, 2000). The *in vitro* finding of ordered assembly of keratin macrofibrils by filaggrin is supported by the effects of filaggrin expression in cultured cells, where overexpression of the monomer leads to the collapse of keratin and vimentin filament networks (Dale *et al.*, 1997; Kuechle *et al.*, 1999) and impairs cell–cell adhesion (Presland *et al.*, 2001). In contrast, profilaggrin is not able to aggregate keratin filaments (Lonsdale-Eccles *et al.*, 1982).

Filaggrin is further processed into free amino acids as terminal differentiation proceeds (Scott and Harding, 1986), leaving only a small amount as a component of the cornified envelope (Steinert and Marekov, 1995; Simon *et al.*, 1996). Of the free amino acids, arginine is deiminated enzymatically (by peptidyl arginine deiminase; Tarcsa *et al.*, 1996), whereas glutamine is modified to pyrrolidone carboxylic acid (Thulin and Walsh, 1995), a highly hydroscopic compound. Although somewhat controversial, the free and modified amino acids have been proposed as an important factor in maintaining stratum corneum moisturization (reviewed in Rawlings and Harding, 2004).

2. *Trichohyalin*

Trichohyalin is a 248-kDa highly charged protein that is predominantly expressed in specialized epithelia, such as the inner root sheeth (IRS) of the hair follicle, which are tailored for resisting mechanical stress. Trichohyalin is also present at lower levels in the filiform papillae of the tongue and some cells of the granular layer of epidermis (Lee *et al.*, 1993; O'Keefe *et al.*, 1993; Presland and Dale, 2000; Rothnagel and Rogers,

1986). Trichohyalin contains extensive repeat sequence blocks, which are predicted to form an elongated, flexible single-stranded α-helical rod (Lee *et al.*, 1993). These predictions are consistent with biochemical and electron microscopical analysis showing an extended shape with overall dimensions of 85 nm. Similar to profilaggrin, the N-terminus comprises an S-100 protein-like calcium-binding domain.

Initially, trichohyalin accumulates in cytoplasmic granules in IRS cells, and can colocalize with filaggrin in granules in some cases (Ishida-Yamamoto *et al.*, 1997; Manabe and O'Guin, 1994; O'Keefe *et al.*, 1993). As differentiation proceeds, trichohyalin associates with IFs (O'Guin *et al.*, 1992; Rothnagel and Rogers, 1986), a redistribution that may depend on posttranslational modification of arginyl residues to citrulline by peptidylarginine deiminase, which renders the protein more soluble and rigid *in vitro* (Tarcsa *et al.*, 1996, 1997).

Trichohyalin is thought to contribute to the scaffolding on which the cell envelope is organized and to anchoring keratin IFs to the cell envelope in the IRS. Extensive crosslinking occurs via the action of transglutaminases with the formation of N^{ε}-(γ-glutamyl)lysine isopeptide crosslinks (Harding and Rogers, 1971, 1972) between trichohyalin and the proteins of the cornified envelope in the IRS—including epiplakin and small proline-rich proteins (Steinert *et al.*, 2003) —and keratin (Dale *et al.*, 1978; Steinert *et al.*, 2003). It has been proposed that the latter connection is facilitated by ionic interactions between trichohyalin and keratin IFs, as the covalent bonds form specifically at two domains (6 and 8) of trichohyalin and the head and tail domains of keratin (Steinert *et al.*, 2003). Thus, in IRS cells, trichohyalin serves not only as a crosslinking component of the cornified envelope and keratin IFs, but also as a linker between the two structures (Steinert *et al.*, 2003). In contrast, in the medulla of the hair, keratin filaments are virtually absent (Powell and Rogers, 1997). There, synthesis of trichohyalin granules continues until they undergo citrulline conversion and transglutaminase crosslinking, finally fusing to form a solid mass (Harding and Rogers, 1976; Powell and Rogers, 1997).

3. Keratin Associated Proteins

Keratin associated proteins (KAPs) represent a large group of proteins that are expressed upon terminal differentiation of hair cells (Rogers and Powell, 1993). KAPs are grouped in three clusters consisting of over 80 proteins in 23 families (Rogers *et al.*, 2002, 2004). KAP molecular weights are predominantly in the 15 to 25 kDa range, but can be as low as 5 kDa and as high as 53 kDa. The families are categorized into high and ultra-high sulphur families with cysteine contents of less than 30 mol% and up

to 36.8 mol%, respectively, and high glycine-tyrosine families with glycine-tyrosine residues making up to 67.8 mol% (KAP20.1) of the amino acid composition (Rogers *et al.*, 2002). The amino acid sequences of the members of most protein families consist of unique repetitive stretches with repeat unit lengths ranging from 6 to 48 amino acids (Rogers and Powell, 1993). The onset of KAP expression generally follows that of keratin IF and is temporarily and spatially regulated in a protein family-specific manner.

In macrofibrils of the hair, KAPs comprise the amorphous matrix that is believed to crosslink keratin bundles in the mature hair (Powell and Rogers, 1997). How KAPs contribute to this crosslinking phenomenon remains controversial. Disulphide bonds between keratin tails and matrix proteins, as well as within and in between different matrix proteins, may play a role. Alternatively, apolar interactions between glycine loops in glycine-tyrosine-rich KAPs type II keratins may take place (Powell and Rogers, 1997).

Intriguingly, KAP-related genes pmg-1 and pmg-2 have been identified in other epidermal appendages including mammary, sebaceous, and sweat glands, in addition to growing hair follicles in skin (Kuhn *et al.*, 1999). Thus, KAPs may play a larger role in the differentiation of multiple epithelial appendage cell types.

B. Proteins that Coordinate with Actin

Several proteins that were originally identified as actin-binding proteins have recently been uncovered as having dual functions in regulating the organization or assembly state of type III IFs. A common theme in these studies is the role of calponin homology (CH) domains—already discussed above as having a prominent role in plakin family structure—in binding to IFs. These studies raise the possibility that CH domains enhance the cytolinker status of plakins by combining both actin- and IF-binding properties within a single domain. The following section discusses other actin-binding proteins that exhibit IF-binding properties.

1. Calponin

Calponin is a widely distributed actin-binding protein thought to play a role in smooth-muscle contraction by inhibiting actomyosin ATPase activity (Small and Gimona, 1998). Its association with dense bodies of smooth muscle cells and widespread presence in other tissues led investigators to speculate that it may play additional roles in cytoskeletal organization (Nishida *et al.*, 1990; Wang and Gusev, 1996). Dense bodies in smooth muscle cells are primary attachment sites for desmin IF (Cooke, 1976).

Consistent with a more general importance in organizing cytoskeletal networks, calponin colocalizes (Mabuchi *et al.*, 1996; North *et al.*, 1994) and copurifies with desmin from dense bodies of smooth muscle cells (Mabuchi *et al.*, 1997). It is also present in the membrane skeleton of smooth muscle, as discussed above. Calponin coprecipitates with (Wang and Gusev, 1996), and possibly bundles, desmin (Nakagawa *et al.*, 1993), and is capable of incorporating into desmin filaments *in vitro* (Mabuchi *et al.*, 1997). However, the authors observed that at low ionic strength, calponin binds with higher affinity to tetrameric desmin than filaments. The likely IF binding site is located in the N-terminal 22 kDa fragment (Wang and Gusev, 1996) containing a CH domain. Collectively, these data raise the possibility that calponin may function to bridge IFs with actin in dense bodies.

2. *Fimbrin*

Fimbrin or plastin is an actin-bundling protein (Bretscher, 1981; Glenney *et al.*, 1981; Namba *et al.*, 1992) consisting of an N-terminal calcium-binding domain followed by four calponin-homology domains. Of the three isoforms, T-fimbrin/plastin has a broad tissue distribution, whereas the I- and L-isoforms are restricted to the intestine, and hematopoietic cells and certain carcinomas, respectively (Lin *et al.*, 1993, 1994). L-fimbrin/plastin was found to coimmunoprecipitate with vimentin from a detergent extract of adherent macrophages, and the N-terminus of vimentin bound to a 45-residue fragment of the CH1-domain in overlay assays (Correia *et al.*, 1999). Fimbrin/plastin most likely binds to tetrameric rather than filamentous vimentin. The *in vivo* interaction, observed by colocalization in immunostained macrophages, occurred in podosomes and filopodia during early adhesion events; it was proposed that fimbrin may direct vimentin assembly at these adhesion sites.

3. *Nebulin*

Ultrastructural studies have long suggested that desmin IFs are linked in some way to the Z-line in striated muscle; however, the molecular basis of this association has not been well defined. A good candidate for performing this function is a protein called nebulin, a component of the thin filament in skeletal and cardiac striated muscle (McElhinny *et al.*, 2003). Nebulin spans from the pointed end of the thin filament where it binds to tropomodulin (McElhinny *et al.*, 2001), to the Z-disk where it binds to actin, tropomyosin, and troponins. Thus, it is speculated that nebulin functions both as a molecular ruler regulating thin filament length (Kruger *et al.*, 1991) and as a scaffold for sarcomeric

thin filaments (McElhinny *et al.*, 2003). However, recent data also suggest that nebulin may have additional functions, ranging from signal transduction to myofibril force generation.

Nebulin is divergently spliced to give rise to a number isoforms with molecular masses of 750–850 kDa, which may provide different templates to tailor sarcomere structure for various physiological functions (Kazmierski *et al.*, 2003). Nebulin's N-terminus associates with the thin filament, and its C-terminus is anchored at the Z-disk (see Fig. 2 in McElhinny *et al.*, 2003). It is the C-terminal region—specifically modules referred to as M163-M170—that localizes to the Z-disk and binds to the 1B helix of the desmin rod (Bang *et al.*, 2002). As desmin laterally links Z-disks together, this interaction is speculated to integrate myofibrils with the sarcolemma and other components of muscle cells (McElhinny *et al.*, 2003) and to maintain Z-lines in register (Bang *et al.*, 2002). The potential importance of this interaction is highlighted by the existence of patients with truncation mutations of the nebulin gene that cause (recessive) nemaline myopathy (Pelin *et al.*, 1999).

V. IFAPs as Participants in a Signaling Scaffold and Cellular Metabolism

In addition to structural roles, it is becoming increasingly apparent that a major function of IFs is to regulate the activities of cell metabolism, by serving as scaffolding for the dynamic regulation of associated proteins such as enzymes, signaling adapter molecules, stress proteins, cell death receptors, and even the endocytic machinery.

A. Stress Proteins/Chaperones

The role of the IF, and particularly the keratin filament system, in resisting the forces of mechanical stress has been well established. However, IFs also play a role in countering metabolic stress. Perhaps the best example is the cytoprotective role played by the simple epithelial keratins, K8/18. However, vimentin, desmin, peripherin, GFAP, the lens proteins phakinin, and filensin and other keratins have also been shown to associate with members of the small heat shock protein (HSP) family, including HSP27 and αB-crystallin (reviewed in Coulombe and Wong, 2004; Marceau *et al.*, 2001; Nicholl and Quinlan, 1994).

Subjecting cells to various forms of stress has been shown to dramatically reorganize IF networks, concomitant with the redistribution of normally soluble αB-crystallins to these networks, resulting in resistance to

detergent extraction. Based on these and similar observations, it has been suggested that αB-crystallin may function as a molecular chaperone for IF proteins (Djabali *et al.*, 1997, 1999). In addition, *in vitro* biochemical analyses have suggested that the association of IFs with chaperones may regulate the organization of filaments by protecting them from non-covalent interactions that could lead to IF aggregation (Perng *et al.*, 1999a). Such a function might explain why mutations in αB-crystallin could underlie the abnormal aggregation seen in certain myopathies and in lens cataract formation (Perng *et al.*, 1999b; Vicart *et al.*, 1998). Keratin 8/18 also associates with other stress-related chaperone-like molecules, including HSP/c70 (Liao *et al.*, 1995; Napolitano *et al.*, 1985), the human 78-kDa glucose-regulated protein (grp78) (Liao *et al.*, 1997), and more recently, Mrj, a DnaJ/HSP40 family protein (Izawa *et al.*, 2000). As Mrj binds directly binds to K18 and HSP/c70, it was proposed that it may act as an adapter to link K8/18 to the HSP/c70 chaperone complex. The fact that microinjection of anti-Mrj antibody disrupted K8/18 networks further supports the contention that these chaperone complexes play an important role in regulating keratin organization. The importance of IFs in protecting the liver from metabolic stress is reflected by the fact that mutations in K8/18 predispose humans to a variety of chronic liver diseases (Ku *et al.*, 1996, 2001, 2003a; Owens *et al.*, 2004; Zatloukal *et al.*, 2000).

B. *Proteins that Regulate Apoptosis*

Among the more recently recognized IFAPs are enzymes and receptors involved in the apoptotic pathway (reviewed in Coulombe and Wong, 2004; Oshima, 2002). Nuclear lamins (Lazebnik *et al.*, 1995), type I keratins (Caulin *et al.*, 1997), vimentin (Byun *et al.*, 2001), and desmin (Chen *et al.*, 2003) are all caspase substrates, and evidence suggests that caspase-generated IF fragments can promote the apoptotic cascade (Byun *et al.*, 2001). Furthermore, recent work has shown that adapter elements of the apoptotic machinery, ubiquitinated DEDD and DEDD2, form a scaffolding dependent on K8/18 and active caspase 3 (Dinsdale *et al.*, 2004; Lee *et al.*, 2002). This observation raises the possibility that IFs might participate in optimally positioning downstream effectors in the cell death pathway.

In addition to these downstream enzymatic effectors, IFs interact with receptors in the death pathway. K8 and K18 interact directly with the cytoplasmic domain of tumor necrosis factor receptor 2 (TNFR2) and attenuate TNF-induced, Jun NH(2)-terminal kinase (JNK) signaling, NFκB activation, and cell death (Caulin *et al.*, 2000). Modulation of cell

surface targeting of Fas has also been observed in cultured cells lacking K8/K18, the loss of which sensitizes cells to Fas-mediated apoptosis (Gilbert *et al.*, 2001), in part by regulating levels of c-Flip, a protein that governs the antiapoptotic ERK1/2 signaling pathway in these cells (Gilbert *et al.*, 2004). The observation that c-Flip physically interacts with K8 and Raf-1 kinase—the latter of which is required for Fas activation of the ERK1/2 pathway—supports the idea that these proteins participate in an antiapoptotic complex. These findings are also in line with the observed sensitivity of transgenic mice with a keratin 18 mutation to Fas-mediated apoptosis (Ku *et al.*, 2003b). A role for K8/18 in governing the targeting of cell surface proteins to their correct domains is supported by studies demonstrating altered distribution of cell surface proteins in intestine and liver of CK8-deficient animals (Ameen *et al.*, 2001b). Recently a direct interaction between both K14 and K18 and the TNFR1-associated death domain protein (TRADD) was observed. The authors showed evidence that keratins can compete with ligand-activated TNFR1 for binding with TRADD (Inada *et al.*, 2001). These data provide an appealing explanation for how the presence of keratins could attenuate cell death, as keratin binding to TRADD would interfere with caspase 8 activation.

C. Interaction of IF with Kinases and Cell Cycle Machinery

Several lines of evidence suggest that IFs serve as scaffolding for intracellular kinases, which not only regulate the phosphorylation and assembly state of IFs, but also govern kinase activity and, in some cases, cell proliferation and differentiation.

Cdk5 has long been known to be an important modulator of the neurofilament network, which is highly sensitive to regulation by phosphorylation (Grant *et al.*, 2001). As its name implies, cdk5 is a cyclin-dependent kinase; however, it has not been found to be involved in cell cycle regulation per se, but is instead primarily active in postmitotic cells such as neurons. Phosphorylation of NF proteins regulates their assembly state, association with other cytoskeletal elements, and transport within neurons. Yeast two hybrid analysis has suggested that the Cdk5 may associate with NF-H via its activator p35, which associates directly with the C-terminal region of the NF-M rod domain (Qi *et al.*, 1998). More recently, another type IV IF protein, nestin, was identified as a novel *in vivo* target for cdk5/p35 kinase. The physical association between nestin and the kinase complex is regulated by the differentiation state of myoblasts. The authors suggest there exists a continuous turnover of cdk5 and p35 activity on a nestin scaffold that affects the organization and function of both cdk5 and nestin during development (Sahlgren *et al.*, 2003).

IFs of the nervous system are not the only ones that serve as binding sites for kinases, which in turn regulate their assembly state. K8/18 have recently been shown to associate with mitogen-activated protein kinase p38; K8 is specifically phosphorylated at a single site, ser-73, in a physiologic manner, causing keratin IF reorganization (Ku *et al.*, 2002a). Introduction of human disease causing mutation in close proximity to this residue increases phosphorylation, suggesting that keratin hyperphosphorylation may contribute to aberrant IF organization and disease pathogenesis.

Association of differentiation-specific keratins with cell cycle machinery also influences proliferation in the epidermis. Paramio and colleagues first demonstrated that K10 inhibits, while K16 promotes, the proliferation of human keratinocytes by acting in some manner on the retinoblastoma (Rb) pathway (Paramio *et al.*, 1999). They later demonstrated that this occurs through a physical interaction of K10 with Akt (protein kinase B [PKB]) and atypical PKCζ. Sequestration of these kinases inhibits their intracellular translocation and interferes with their activation, consequently impairing pRb phosphorylation, reducing the expression of cyclins D1 and E (Paramio *et al.*, 2001) and reducing NFκB activity (Santos *et al.*, 2003). Consistent with the observed ability of K10 to interfere with these cell cycle pathways are *in vivo* studies showing that mice ectopically expressing K10 in the proliferating basal layers of the epidermis exhibit a hypoplastic and hyperkeratotic epidermis. The association of IFs with adapter proteins such as 14-3-3 also has a potential impact on regulation of the cell cycle machinery, as detailed in the section that follows.

D. 14-3-3 Proteins

14-3-3 proteins are adapters that interact with specific phosphoserine or threonine motifs in a plethora of cell cycle proteins, transcription factors, and other binding partners involved in signal transduction cascades (Bridges and Moorhead, 2004). It is thought that binding of 14-3-3 proteins to their partners either changes protein conformation, physically masks certain features, or serves a scaffolding function (Bridges and Moorhead, 2004). The protein 14-3-3 has been shown to interact directly with both vimentin and keratin IF. Cell cycle dependent phosphorylation of K18 leads to its association with 14-3-3 (Ku *et al.*, 1998), which in turn modulates keratin filament organization and acts as a solubility cofactor (Liao and Omary, 1996). But 14-3-3 also regulates mitotic progression of cells expressing K8/18, such as hepatocytes (Ku *et al.*, 2002b). In mice lacking K8/18, cell cycle regulation is disturbed and 14-3-3 is redistributed

to nuclei. This change in localization could, in turn, alter the association of 14-3-3 with other important cell cycle regulators (Coulombe and Wong, 2004; Toivola *et al.*, 2001). Also, 14-3-3 interacts with phosphovimentin, displacing the 14-3-3 partner Raf, which can no longer be activated by EGF (Tzivion *et al.*, 2000). In general, it seems that interaction of 14-3-3 with IFs interferes with other effector pathways. In an interesting twist, studies have shown that the 14-3-3 binding partner Raf can also bind directly to K8/18, the latter interaction which blocks keratin–14-3-3 binding in a reconstitution assay (Ku *et al.*, 2004). The authors' results suggest that K8/18 can sequester a unique phosphospecific population of Raf, consequently regulating its signaling functions in a manner that is analogous to 14-3-3's regulation of Raf-1. Whether IFs might also in some cases facilitate 14-3-3-dependent interactions with effectors awaits further study.

E. Interaction with Lipids and Membranes: Involvement in Membrane Trafficking

Reports dating back to the mid 1980s describe a potential direct association between vimentin and membrane lipids (Traub *et al.*, 1985, 1987). Other observations suggested that vimentin may play a novel role in differentiating 3T3 cells during adipose conversion by participating in the compartmentalization of the forming lipid globules (Franke *et al.*, 1987). More recent work utilizing vimentin-deficient cells has demonstrated that the intracellular movement of LDL-derived cholesterol from the lysosome to the site of esterification is impaired (Sarria *et al.*, 1992) and the cells show defects in glycolipid synthesis, possibly due to a defect in transport of glycosphigolipids between the endosomal/lysosomal pathway and the Golgi apparatus. A direct interaction between vimentin and the Golgi complex through the peripherally associated Golgi protein formiminotransferase cyclodeaminase (FTCD) has been demonstrated. Neither the interaction, nor the ability of FTCD to promote filament assembly, require enzyme activity (Gao *et al.*, 2002).

The observation that cell surface marker distribution and transport of components in the apoptotic cascade are altered in K8/18 deficient cells (Ameen *et al.*, 2001a; Gilbert *et al.*, 2001) raises the possibility that IFs play a more general role in the organization of intracellular membranes and trafficking. Supporting this idea, regional specific differences in the distribution of syntaxin-3, a central player in the SNARE machinery involved in vesicular transport, were observed in the intestines of K8 null cells. Furthermore, the SNARE SNAP23 associates with vimentin filaments in primary fibroblasts and HeLa cells (Faigle *et al.*, 2000).

Very recently, a novel association between rat brain peripherin and the AP-3, an adapter protein involved in the formation of lysosomes, lysosome-related organelles, and synaptic vesicles, was identified. The authors extended their work to show that other type III IFs, including α-internexin and vimentin, also associate with AP-3 (Styers *et al.*, 2004). The authors went on to show that in vimentin-deficient fibroblasts, AP-3 and late endosomal/lysosomal markers are redistributed and AP-3-dependent sorting is impaired. Together, these data suggest that the IF cytoskeleton participates in regulating membrane architecture and compartmentalization, as well as trafficking events in the endolysosomal pathway.

VI. Motors that Move IFs

Some IFAPs are motor molecules that play a role in positioning IFs in various tissues and in regulating the dynamic nature of the IF cytoskeleton (reviewed in Chang and Goldman, 2004; Helfand *et al.*, 2004). Far from being the static struts, IFs and their precursors have been shown to be dynamic elements of the cytoskeleton that are whisked around the cell by molecular motors. Some of the first data demonstrating that IFs exist as precursor particles that are moved around the cell came from studies in the Goldman lab. This group demonstrated that GFP-vimentin incorporates into small vimentin particles in BHK fibroblasts that colocalize with the anterograde MT motor kinesin and move rapidly in a kinesin-dependent manner (Yoon *et al.*, 1998). The kinesin-heavy and specific 62kDa light chain appear to be required for mediating interactions with vimentin (Avsiuk *et al.*, 1995; Liao and Gundersen, 1998). The particles eventually coalesce into larger structures referred to as *squiggles*, which in turn are postulated to assemble into longer filaments. Although anterograde movements are more common, movements can occur in both directions. More recently, the same group demonstrated that the dynein/dynactin complex is also associated with vimentin and is responsible for retrograde movement of particles, squiggles, and filaments (Helfand *et al.*, 2002). Thus, a balance of plus and minus end movements is thought to be important for the homeostasis of the intermediate filament network.

Vimentin is not the only IF type that is moved in this manner. Neural IFs, including both the type III peripherin proteins and type IV neural IF proteins, have also been demonstrated to be moved by molecular motors. In the case of peripherin, particles and squiggles were observed to translocate rapidly within PC12 cell bodies, neurites, and growth cones. The movements were bidirectional and dependent on microtubules, kinesin, and cytoplasmic dynein. The authors suggest that peripherin particles are

part of a rapid transport system that delivers cytoskeletal subunits to distal regions of neurites (Helfand *et al.*, 2003).

Studies have demonstrated that GFP-tagged NF-M and NF-H form particles and move in an MT-dependent fashion, similar to that demonstrated for peripherin (reviewed in Prahlad *et al.*, 2000; Shah *et al.*, 2000; Shea, 2000; Yabe *et al.*, 1999). These studies also help to explain the underlying basis of slow axonal transport. NF particles and NF up to about 16 μm actually move at high rates, but their transport is punctuated by long pauses. The result is net movement consistent with slow axonal transport (Wang *et al.*, 2000). Intriguingly, peripherin pauses two to three times less frequently than NF, in spite of similar MT-dependent mechanisms of transport. Differences in IF structure—such as the existence of the long, highly charged tails on NF-M and NF-H that might act as cross-bridges to stop motility—and regulation of molecular motor association by phosphorylation could also account for these differences (Yabe *et al.*, 2000).

Myosin Va is also associated with NFs, with NF-L being the major binding partner (Rao *et al.*, 2002). This association has been shown to be important in regulating the localization and content of a large pool of myosin Va in neurons, in addition to regulating NF number. Association with the actomyosin system provides another potential transport mechanism and an opportunity for more sophisticated modes of regulation within the peripheral nervous system.

Nudel, first identified as a protein required for nuclear migration in filamentous fungi, was identified as a critical NF associated protein in mammals. Its loss results in disruption of NF organization and neurodegenerative changes (Holzbaur, 2004; Nguyen *et al.*, 2004). The results suggest that Nudel is important for assembling NF cytoskeleton through its direct interaction with the central rod domains of NF-L. As nudel also interacts with dynein, it also seems possible that its role in assembly is coupled with MT-dependent transport of NF down the axon. Interfering with this process could potentially lead to NF accumulation and signs of neurodegeneration.

Mutations in motor proteins or IFs themselves (which may alter their associations with IFAPs or motors) lead to accumulations of IFs in ALS, Charcot-Marie Tooth disease 2, and Parkinson's (Goldstein and Yang, 2000; Helfand *et al.*, 2004). Impaired assembly and transport of NFs is a critical determinant of neurodegenerative disease. Consistent with a critical role for kinesin *in vivo*, mice lacking the neuronal-specific conventional kinesin heavy chain KIF5A were shown to have accumulations of NF-H, as well as NF-M and NF-L, in the cell bodies of peripheral sensory neurons. The presence of these accumulations was accompanied by a reduction in

sensory axon caliber. These data support the hypothesis that a conventional kinesin plays a role in the microtubule-dependent slow axonal transport of at least one cargo, the NF proteins (Xia *et al.*, 2003).

Keratin IFs exhibit different motile behaviors compared with type III and IV IFs, and continue to move in the absence of MT in cells treated with MT inhibitors, such as nocodazole. It has been proposed that actin may be involved in their transport (Yoon *et al.*, 2001). Time lapse imaging of keratin-GFP has revealed that keratin IF formation begins in the cell cortex in a region enriched in actin and required for both intact MT and actin, and that actin governs the organization and movement of keratin in Xenopus egg extracts (Weber and Bement, 2002). Less is known about specific mediators of these interactions, although the fact that myosin Va associates with NFs raises the interesting possibility that other such interactions may exist for keratins.

Collectively, these observations suggest that multiple interactions between motor proteins and various cell-type-specific IFs have evolved to correctly deploy IF precursors in the cytoplasm, where they can appropriately polymerize and perform their functions without interfering with other aspects of cellular metabolism. These precursors can take on a variety of forms, from particles to short filaments. It has been demonstrated that different-sized IFs move at different speeds, but it seems unlikely that these differences in rates of movement are directly related to size, as kinesin moves cargoes equally irrespective of size. It has been proposed that IFAPs, such as crosslinkers in the plakin family, may tether longer structures to cytoarchitecture, thus conferring another level of regulation to the process of IF transport and assembly (Chang and Goldman, 2004).

VII. The Fine Line between IFs and IFAPs

The classification of proteins as IFs has been debated over the years, with much discussion as to whether certain IFs are *bona fide* IFs or IFAPs; then, if they are IFs, how should they be classified? In the end, the classification has been based primarily on sequence homology and protein structure. Thus, synemin and paranemin, which were originally identified as IFAPs, have now been dubbed IFs (Bellin *et al.*, 1999; Hemken *et al.*, 1997). In spite of sequence similarities, however, they differ from classical IFs in that these proteins require other IFs (such as desmin) for incorporation into the polymer. In this respect, synemin and paranemin are similar to several other IFs including nestin, NF-H, tanabin in the nervous system (for review, see Coulombe *et al.*, 2000) and syncoilin in muscle (Blake and Martin-Rendon, 2002).

Another feature of these proteins is the divergence of their nonhelical domains, which are sometimes quite extended and play important roles in organizing the filaments with which they copolymerize. For instance, the 661-residue-long NF-H tail projects away from the IF surface, providing a crosslinking function important for interfilament spacing in neurons (Lee and Cleveland, 1996). The incorporation of substoichiometric amounts of nestin into type III and IV IFs may play a similar role in spacing these filaments within cells (Steinert *et al.*, 1999). Likewise, the tail domain of synemin provides a binding site for both desmin and α-actinin, thus integrating IF and actin systems at costomeres and z-lines in striated muscle (Bellin *et al.*, 2001). Together, these data suggest that IF networks employ multiple mechanisms to regulate their organization in different cells and tissues, sometimes incorporating spacer IFAPs into the core of the polymer, and other times, recruiting peripheral proteins from which the polymer may more readily dissociate during remodeling events.

VIII. Conclusions and Future Directions

The number of IF genes has risen to more than 67, and IF gene expression is exquisitely regulated during development and differentiation. As the IF family has broadened, so too has our definition of what constitutes an IFAP. It is now clear that IFAPs include among their number more than cytolinkers and bundling proteins. Enzymes, molecular motors, membrane receptors, adapters, and other molecules important for cell homeostasis and metabolism have all been identified as directly binding to IFs and modifying their functions. The varied and regulated associations between IFs and their partners dramatically increase the functional complexity of the cytoplasmic IF scaffold.

There is much more to this functional complexity than we have been able to discuss in this review. For instance, we have not touched on the complex intranuclear network mediated by the nuclear lamins, which is critically important for human health as reflected in the numerous disease phenotypes that are emerging due to mutation in lamins and their associated protein partners. We have also not discussed the association of IFs with molecules involved in bacterial and viral pathogenesis, such as the Ebstein-Barr virus latent infection membrane protein (LMP) (Liebowitz *et al.*, 1987) or viral kinases such as the herpes simplex virus US3 kinase (Murata *et al.*, 2002) that may be important in remodeling IFs during infection. It is clear, however, from the many critical cellular functions we have discussed in this review, that the IF structural and signaling scaffold is involved in wide-ranging functions that are important for fundamental

cellular processes, and that the diversity of the IF family and their partnerships allows for tailoring these functions to the many cell types and tissues found in higher organisms.

Acknowledgments

The authors would like to thank Spiro Getsios for critical reading of the manuscript and Ron Liem and Bishr Omary for helpful discussions, and also our colleagues who shared information prior to its publication. Work in the authors' laboratories is supported by NIH grants RO1 DK60589 , RO1 HL067016, PO1 HL071643, PO1 DE12328 (project #3) to J. Jones and RO1AR41836, RO1 AR43380, and PO1 DE12328 (project #4) to K. Green.

References

Aho, S., Li, K., Ryoo, Y., McGee, C., Ishida-Yamamoto, A., Uitto, J., and Klement, J. F. (2004). Periplakin gene targeting reveals a constituent of the cornified cell envelope dispensable for normal mouse development. *Mol. Cell. Biol.* **24,** 6410–6418.

Alcalai, R., Metzger, S., Rosenheck, S., Meiner, V., and Chajek-Shaul, T. (2003). A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder, and woolly hair. *J. Am. Coll. Cardiol.* **42,** 319–327.

Ameen, N. A., Figueroa, Y., and Salas, P. J. (2001a). Anomalous apical plasma membrane phenotype in CK8-deficient mice indicates a novel role for intermediate filaments in the polarization of simple epithelia. *J. Cell Sci.* **114,** 563–575.

Ameen, N. A., Figueroa, Y., and Salas, P. J. (2001b). Anomalous apical plasma membrane phenotype in CK8-deficient mice indicates a novel role for intermediate filaments in the polarization of simple epithelia. *J. Cell Sci.* **114,** 563–575.

Anastasiadis, P. Z., and Reynolds, A. B. (2000). The p120 catenin family: Complex roles in adhesion, signaling and cancer. *J. Cell Sci.* **113,** 1319–1334.

Andrä, K., Kornacker, I., Jörgl, A., Zörer, M., Spazierer, D., Fuchs, P., Fischer, I., and Wiche, G. (2003). Plectin-isoform-specific rescue of hemidesmosomal defects in plectin (-/-) keratinocytes. *J. Invest. Dermatol.* **120,** 189–197.

Andrä, K., Lassmann, H., Bittner, R., Shorny, S., Fässler, R., Propst, F., and Wiche, G. (1997). Targeted inactivation of plectin reveals essential function in maintaining the integrity of skin, muscle, and heart cytoarchitecture. *Genes Dev.* **11,** 3143–3156.

Andrä, K., Nikolic, B., Stöcher, M., Drenckhahn, D., and Wiche, G. (1998). Not just scaffolding: Plectin regulates actin dynamics in cultured cells. *Genes Dev.* **12,** 3442–3451.

Angst, B. D., Nilles, L. A., and Green, K. J. (1990). Desmoplakin II expression is not restricted to stratified epithelia. *J. Cell Sci.* **97,** 247–257.

Armstrong, D. K. B., McKenna, K. E., Purkis, P. E., Green, K. J., Eady, R. A. J., Leigh, I. M., and Hughes, A. E. (1999). Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. *Hum. Mol. Genet.* **8,** 143–148.

Avsiuk, A., Minin, A., and Gyoeva, F. (1995). Kinesin associated wth vimentin intermediate filaments contains a specific light chain. *Doklady Biol. Sci.* **345,** 644–646.

Bang, M. L., Gregorio, C., and Labeit, S. (2002). Molecular dissection of the interaction of desmin with the C-terminal region of nebulin. *J. Struct. Biol.* **137,** 119–127.

Bellin, R. M., Huiatt, T. W., Critchley, D. R., and Robson, R. M. (2001). Synemin may function to directly link muscle cell intermediate filaments to both myofibrillar Z-lines and costameres. *J. Biol. Chem.* **276,** 32330–32337.

Bellin, R. M., Sernett, S. W., Becker, B., Ip, W., Huiatt, T. W., and Robson, R. M. (1999). Molecular characteristics and interactions of the intermediate filament protein synemin. Interactions with alpha-actinin may anchor synemin-containing heterofilaments. *J. Biol. Chem.* **274,** 29493–29499.

Bershadsky, A. D., Tint, I. S., and Svitkina, T. M. (1987). Association of intermediate filaments with vinculin-containing adhesion plaques of fibroblasts. *Cell Motil. Cytoskel.* **8,** 274–283.

Blake, D. J., and Martin-Rendon, E. (2002). Intermediate filaments and the function of the dystrophin-protein complex. *Trends Cardiovasc. Med.* **12,** 224–228.

Bonne, S., Gilbert, B., Hatzfeld, M., Chen, X., Green, K. J., and van Roy, F. (2003). Defining desmosomal plakophilin-3 interactions. *J. Cell Biol.* **161,** 403–416.

Bornslaeger, E., Godsel, L., Corcoran, C., Park, J., Hatzfeld, M., Kowalczyk, A., and Green, K. (2001). Plakophilin 1 interferes with plakoglobin binding to desmoplakin, yet together with plakoglobin promotes clustering of desmosomal plaque complexes at cell–cell borders. *J. Cell Sci.* **114,** 727–738.

Bornslaeger, E. B., Corcoran, C. M., Stappenbeck, T. S., and Green, K. J. (1996). Breaking the connection: Displacement of the desmosomal plaque protein desmoplakin from cell–cell interfaces disrupts anchorage of intermediate filament bundles and alters intercellular junction assembly. *J. Biol. Chem.* **134,** 985–1002.

Borradori, L., and Sonnenberg, A. (1999). Structure and function of hemidesmosomes: More than simple adhesion complexes. *J. Invest. Derm.* **112,** 411–418.

Brakebusch, C., and Fassler, R. (2003). The integrin-actin connection, an eternal love affair. *Embo. J.* **22,** 2324–2333.

Bretscher, A. (1981). Fimbrin is a cytoskeletal protein that crosslinks F-actin *in vitro*. *Proc. Natl. Acad. Sci. USA* **78,** 6849–6853.

Bridges, D., and Moorhead, G. B. (2004). 14-3-3 proteins: A number of functions for a numbered protein. *Sci. STKE* **2004**(242), re10.

Brown, A., Bernier, G., Mathieu, M., Rossant, J., and Kothary, R. (1995). The mouse dystonia musculorum gene is a neural isoform of bullous pemphigoid antigen 1. *Nature Genet.* **10,** 301–306.

Byun, Y., Chen, F., Chang, R., Trivedi, M., Green, K. J., and Cryns, V. L. (2001). Caspase cleavage of vimentin disrupts intermediate filaments and promotes apoptosis. *Cell Death Differ.* **8,** 443–450.

Calkins, C. C., Hoepner, B. L., Law, C. M., Novak, M. R., Setzer, S. V., Hatzfeld, M., and Kowalczyk, A. P. (2003). The Armadillo family protein p0071 is a VE-cadherin- and desmoplakin-binding protein. *J. Biol. Chem.* **278,** 1774–1783.

Carlsson, L., and Thornell, L. E. (2001). Desmin-related myopathies in mice and man. *Acta Physiol. Scand.* **171,** 341–348.

Cartaud, A., Ludosky, M. A., Courvalin, J. C., and Cartaud, J. (1990). A protein antigenically related to nuclear lamin B mediates the association of intermediate filaments with desmosomes. *J. Biol. Chem.* **111,** 581–588.

Caulin, C., Salvesen, G. S., and Oshima, R. G. (1997). Caspase cleavage of keratin 18 and reorganization of intermediate filaments during epithelial cell apoptosis. *J. Cell Biol.* **138,** 1379–1394.

Caulin, C., Ware, C. F., Magin, T. M., and Oshima, R. G. (2000). Keratin-dependent, epithelial resistance to tumor necrosis factor-induced apoptosis. *J. Cell Biol.* **149,** 17–22.

Chang, L., and Goldman, R. (2004). Intermediate filaments mediate cytoskeletal crosstalk. *Nat. Rev. Mol. Cell. Biol.* **5,** 601–613.

Chen, F., Chang, R., Trivedi, M., Capetanaki, Y., and Cryns, V. L. (2003). Caspase proteolysis of desmin produces a dominant-negative inhibitor of intermediate filaments and promotes apoptosis. *J. Biol. Chem.* **278,** 6848–6853.

Chen, X., Bonne, S., Hatzfeld, M., van Roy, F., and Green, K. J. (2002). Protein binding and functional characterization of plakophilin 2. Evidence for its diverse roles in desmosomes and beta-catenin signaling. *J. Biol. Chem.* **277,** 10512–10522.

Choi, H. J., Park-Snyder, S., Pascoe, L. T., Green, K. J., and Weis, W. I. (2002). Structures of two intermediate filament-binding fragments of desmoplakin reveal a unique repeat motif structure. *Nat. Struct. Biol.* **9,** 612–620.

Clubb, B. H., Chou, Y. H., Herrmann, H., Svitkina, T. M., Borisy, G. G., and Goldman, R. D. (2000). The 300-kDa intermediate filament-associated protein (IFAP300) is a hamster plectin ortholog. *Biochem. Biophys. Res. Commun.* **273,** 183–187.

Cooke, P. (1976). A filamentous cytoskeleton in vertebrate smooth muscle fibers. *J. Biol. Chem.* **68,** 539–556.

Correia, I., Chu, D., Chou, Y. H., Goldman, R. D., and Matsudaira, P. (1999). Integrating the actin and vimentin cytoskeletons: Adhesion-dependent formation of fimbrin-vimentin complexes in macrophages. *J. Cell Biol.* **146,** 831–842.

Coulombe, P. A., Bousquet, O., Ma, L., Yamada, S., and Wirtz, D. (2000). The 'ins' and 'outs' of intermediate filament organization. *Trends Cell Biol.* **10,** 420–428.

Coulombe, P. A., and Omary, M. B. (2002). 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Curr. Opin. Cell Biol.* **14,** 110–122.

Coulombe, P. A., and Wong, P. (2004). Cytoplasmic intermediate filaments revealed as dynamic and multipurpose scaffolds. *Nat. Cell Biol.* **6,** 699–706.

Craig, S. W., and Pardo, J. V. (1983). Gamma actin, spectrin, and intermediate filament proteins colocalize with vinculin at costameres, myofibril-to-sarcolemma attachment sites. *Cell Motil.* **3,** 449–462.

Dale, B. A., Holbrook, K. A., and Steinert, P. M. (1978). Assembly of stratum corneum basic protein and keratin filaments in macrofibrils. *Nature* **276,** 729–731.

Dale, B. A., and Presland, R. B. (1999). Filaggrins. *In* ''Guidebook to the Cytoskeletal and Motor Proteins'' (T. Kreis and R. Vale, Eds.), pp. 333–337. Oxford University Press, New York.

Dale, B. A., Presland, R. B., Fleckman, P., Kam, E., and Resing, K. A. (1993). Phenotypic expression and processing of filaggrin in epidermal differentiation. *In* ''Molecular Biology of the Skin'' (M. Darmon and M. Blumenberg, Eds.), pp. 79–106. Academic Press, San Diego.

Dale, B. A., Presland, R. B., Lewis, S. P., Underwood, R. A., and Fleckman, P. (1997). Transient expression of epidermal filaggrin in cultured cells causes collapse of intermediate filament networks with alteration of cell shape and nuclear integrity. *J. Invest. Dermatol.* **108,** 179–187.

Dale, B. A., Resing, K. A., and Presland, R. B. (1994). Keratohyalin granule proteins. *In* ''The Keratinocyte Handbook'' (I. M. Leigh, E. B. Lane, and F. M. Watt, Eds.), pp. 323–350. Cambridge University Press, Cambridge UK.

Dalpe, G., Mathieu, M., Comtois, A., Zhu, E., Wasiak, S., De Repentigny, Y., Leclerc, N., and Kothary, R. (1999). Dystonin-deficient mice exhibit an intrinsic muscle weakness and an instability of skeletal muscle cytoarchitecture. *Dev. Biol.* **210,** 367–380.

Danowski, B. A., Imanaka-Yoshida, K., Sanger, J. M., and Sanger, J. W. (1992). Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes. *J. Biol. Chem.* **118,** 1411–1420.

DiColandrea, T., Karashima, T., Maatta, A., and Watt, F. M. (2000). Subcellular distribution of envoplakin and periplakin: Insights into their role as precursors of the epidermal cornified envelope. *J. Cell Biol.* **151,** 573–586.

Dinsdale, D., Lee, J. C., Dewson, G., Cohen, G. M., and Peter, M. E. (2004). Intermediate filaments control the intracellular distribution of caspases during apoptosis. *Am. J. Pathol.* **164,** 395–407.

Djabali, K., de Nechaud, B., Landon, F., and Portier, M. M. (1997). AlphaB-crystallin interacts with intermediate filaments in response to stress. *J. Cell Sci.* **110,** 2759–2769.

Djabali, K., Piron, G., de Nechaud, B., and Portier, M. M. (1999). alphaB-crystallin interacts with cytoplasmic intermediate filament bundles during mitosis. *Exp. Cell Res.* **253,** 649–662.

Eckes, B., Colucci-Guyon, E., Smola, H., Nodder, S., Babinet, C., Krieg, T., and Martin, P. (2000). Impaired wound healing in embryonic and adult mice lacking vimentin. *J. Cell Sci.* **113,** 2455–2462.

Eckes, B., Dogic, D., Colucci-Guyon, E., Wang, N., Maniotis, A., Ingber, D., Merckling, A., Langa, F., Aumailley, M., Delouvee, A., Koteliansky, V., Babinet, C., and Krieg, T. (1998). Impaired mechanical stability, migration and contractile capacity in vimentin-deficient fibroblasts. *J. Cell Sci.* **111,** 1897–1907.

Eger, A., Stockinger, A., Wiche, G., and Foisner, R. (1997). Polarisation-dependent association of plectin with desmoplakin and the lateral submembrane skeleton in MDCK cells. *J. Cell Sci.* **110,** 1307–1316.

Elliott, C. E., Becker, B., Oehler, S., Castanon, M. J., Hauptmann, R., and Wiche, G. (1997). Plectin transcript diversity: Identification and tissue distribution of variants with distinct first coding exons and rodless isoforms. *Genomics* **42,** 115–125.

Ervasti, J. M. (2003). Costameres: The Achilles' heel of Herculean muscle. *J. Biol. Chem.* **278,** 13591–13594.

Faigle, W., Colucci-Guyon, E., Louvard, D., Amigorena, S., and Galli, T. (2000). Vimentin filaments in fibroblasts are a reservoir for SNAP23, a component of the membrane fusion machinery. *Mol. Biol. Cell* **11,** 3485–3494.

Fairley, J. A., Scott, G. A., Jensen, K. D., Goldsmith, L. A., and Diaz, L. A. (1991). Characterization of keratocalmin, a calmodulin-binding protein from human epidermis. *J. Clin. Invest.* **88,** 315–322.

Foisner, R., Bohn, W., Mannweiler, K., and Wiche, G. (1995). Distribution and ultrastructure of plectin arrays in subclones of rat glioma C6 cells differing in intermediate filament protein (vimentin) expression. *J. Struct. Biol.* **115,** 304–317.

Foisner, R., Leichtfried, F. E., Herrmann, H., Small, J. V., Lawson, D., and Wiche, G. (1988). Cytoskeleton-associated plectin: *in situ* localization, *in vitro* reconstitution, and binding to immobilized intermediate filament proteins. *J. Biol. Chem.* **106,** 723–733.

Foisner, R., Traub, P., and Wiche, G. (1991). Protein kinase A- and protein kinase C-regulated interaction of plectin with lamin B and vimentin. *Proc. Natl. Acad. Sci.* **88,** 3812–3816.

Foisner, R., and Wiche, G. (1987). Structure and hydrodynamic properties of plectin molecules. *J. Mol. Biol.* **198,** 515–531.

Fontao, L., Favre, B., Riou, S., Geerts, D., Jaunin, F., Saurat, J. H., Green, K. J., Sonnenberg, A., and Borradori, L. (2003). Interaction of the bullous pemphigoid

antigen 1 (BP230) and desmoplakin with intermediate filaments is mediated by distinct sequences within their COOH terminus. *Mol. Biol. Cell* **14,** 1978–1992.

Fontao, L., Geerts, D., Kuikman, I., Koster, J., Kramer, D., and Sonnenberg, A. (2001). The interaction of plectin with actin: Evidence for cross-linking of actin filaments by dimerization of the actin-binding domain of plectin. *J. Cell Sci.* **114,** 2065–2076.

Franke, W. W., Hergt, M., and Grund, C. (1987). Rearrangement of the vimentin cytoskeleton during adipose conversion: Formation of an intermediate filament cage around lipid globules. *Cell* **49,** 131–141.

Frappier, T., Derancourt, J., and Pradel, L.-A. (1992). Actin and neurofilament binding domain of brain spectrin beta subunit. *Eur. J. Biochem.* **205,** 85–91.

Frappier, T., Stetzkowski-Marden, F., and Pradel, L.-A. (1991). Interaction domains of neurofilament light chain and brain spectrin. *Biochem. J.* **275,** 521–527.

Fuchs, E., and Yang, Y. (1999). Crossroads on cytoskeletal pathways. *Cell* **98,** 547–550.

Fuchs, P., Zorer, M., Rezniczek, G. A., Spazierer, D., Oehler, S., Castanon, M. J., Hauptmann, R., and Wiche, G. (1999). Unusual 5′ transcript complexity of plectin isoforms: Novel tissue-specific exons modulate actin binding activity. *Hum. Mol. Genet.* **8,** 2461–2472.

Fujiwara, S., Kohno, K., Iwamatsu, A., Naito, I., and Shinkai, H. (1996). Identification of a 450-kDa human epidermal autoantigen as a new member of the plectin family. *J. Invest. Derm.* **106,** 1125–1130.

Fujiwara, S., Takeo, N., Otani, Y., Parry, D. A., Kunimatsu, M., Lu, R., Sasaki, M., Matsuo, N., Khaleduzzaman, M., and Yoshioka, H. (2001). Epiplakin, a novel member of the Plakin family originally identified as a 450-kDa human epidermal autoantigen. Structure and tissue localization. *J. Biol. Chem.* **276,** 13340–13347.

Gache, Y., Chavanas, S., Lacour, J. P., Wiche, G., Owaribe, K., Meneguzzi, G., and Ortonne, J. P. (1996). Defective expression of plectin/HD1 in epidermolysis bullosa simplex with muscular dystrophy. *J. Clin. Invest.* **97,** 2289–2298.

Gallicano, G. I., Bauer, C., and Fuchs, E. (2001). Rescuing desmoplakin function in extra-embryonic ectoderm reveals the importance of this protein in embryonic heart, neuroepithelium, skin and vasculature. *Development* **128,** 929–941.

Gallicano, G. I., Kouklis, P., Bauer, C., Yin, M., Vasioukhin, V., Degenstein, L., and Fuchs, E. (1998). Desmoplakin is required early in development for assembly of desmosomes and cytoskeletal linkage. *J. Biol. Chem.* **143,** 2009–2022.

Gan, S. Q., McBride, O. W., Idler, W. W., Markova, N., and Steinert, P. M. (1990). Organization, structure, and polymorphisms of the human profilaggrin gene. *Biochemistry* **29,** 9432–9440.

Gao, Y. S., Vrielink, A., MacKenzie, R., and Sztul, E. (2002). A novel type of regulation of the vimentin intermediate filament cytoskeleton by a Golgi protein. *Eur. J. Cell Biol.* **81,** 391–401.

Garcia-Alvarez, B., Bobkov, A., Sonnenberg, A., and de Pereda, J. M. (2003). Structural and functional analysis of the actin binding domain of plectin suggests alternative mechanisms for binding to F-actin and integrin beta4. *Structure (Camb.)* **11,** 615–625.

Gaudry, C. A., Palka, H. L., Dusek, R. L., Huen, A. C., Khandekar, M. J., Hudson, L. G., and Green, K. J. (2001). Tyrosine phosphorylated plakoglobin is associated with desmogleins but not desmoplakin after EGFR activation. *J. Biol. Chem.* **276,** 24871–24880.

Geerts, D., Fontao, L., Nievers, M. G., Schaapveld, R. Q., Purkis, P. E., Wheeler, G. N., Lane, E. B., Leigh, I. M., and Sonnenberg, A. (1999). Binding of integrin

alpha6beta4 to plectin prevents plectin association with F-actin but does not interfere with intermediate filament binding. *J. Cell Biol.* **147,** 417–434.

Georgatos, S. D., and Blobel, G. (1987a). Lamin B constitutes an intermediate filament attachment site at the nuclear envelope. *J. Biol. Chem.* **105,** 117–125.

Georgatos, S. D., and Blobel, G. (1987b). Two distinct attachment sites for vimentin along the plasma membrane and the nuclear envelope in avian erythrocytes: A basis for a vectorial assembly of intermediate filaments. *J. Biol. Chem.* **105,** 105–115.

Georgatos, S. D., and Marchesi, V. T. (1985). The binding of vimentin to human erythrocyte membranes: A model system for the study of intermediate filament-membrane interactions. *J. Biol. Chem.* **100,** 1955–1961.

Georgatos, S. D., Weaver, D. C., and Marchesi, V. T. (1985). Site specificity in vimentin-membrane interactions: Intermediate filament subunits associate with the plasma membrane via their head domains. *J. Biol. Chem.* **100,** 1962–1967.

Georgatos, S. D., Weber, K., Geisler, N., and Blobel, G. (1987). Binding of two desmin derivatives to the plasma membrane and the nuclear envelope of avian erythrocytes: Evidence for a conserved site-specificity in intermediate filament-membrane interactions. *Proc. Natl. Acad. Sci.* **84,** 6780–6784.

Getsios, S., Huen, A. C., and Green, K. J. (2004). Working out the strength and flexibility of desmosomes. *Nat. Rev. Mol. Cell. Biol.* **5,** 271–281.

Gilbert, S., Loranger, A., Daigle, N., and Marceau, N. (2001). Simple epithelium keratins 8 and 18 provide resistance to Fas-mediated apoptosis. The protection occurs through a receptor-targeting modulation. *J. Cell Biol.* **154,** 763–774.

Gilbert, S., Loranger, A., and Marceau, N. (2004). Keratins modulate c-Flip/extracellular signal-regulated kinase 1 and 2 antiapoptotic signaling in simple epithelial cells. *Mol. Cell. Biol.* **24,** 7072–7081.

Giorda, R., Cerritello, A., Bonaglia, M. C., Bova, S., Lanzi, G., Repetti, E., Giglio, S., Baschirotto, C., Pramparo, T., Avolio, L. Bragheri, R., Maraschio, P., and Zuffardi, O. (2004). Selective disruption of muscle and brain-specific BPAG1 isoforms in a girl with a 6;15 translocation, cognitive and motor delay, and tracheo-oesophageal atresia. *J. Med. Genet.* **41,** e71.

Glenney, J. R., Jr., Kaulfus, P., Matsudaira, P., and Weber, K. (1981). F-actin binding and bundling properties of fimbrin, a major cytoskeletal protein of microvillus core filaments. *J. Biol. Chem.* **256,** 9283–9288.

Godsel, L. M., Getsios, S., Huen, A. C., and Green, K. J. (2004). The molecular composition and function of desmosomes. *In* ''Handbook of Experimental Pharmacology'' (K. Starke, Ed.), Vol. 165, pp. 137–193. Springer-Verlag, Berlin.

Goldman, R. D., Gruenbaum, Y., Moir, R. D., Shumaker, D. K., and Spann, T. P. (2002). Nuclear lamins: Building blocks of nuclear architecture. *Genes Dev.* **16,** 533–547.

Goldstein, L. S., and Yang, Z. (2000). Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. *Annu. Rev. Neurosci.* **23,** 39–71.

Gonzales, M., Weksler, B., Tsuruta, D., Goldman, R. D., Yoon, K. J., Hopkinson, S. B., Flitney, F. W., and Jones, J. C. (2001). Structure and function of a vimentin-associated matrix adhesion in endothelial cells. *Mol. Biol. Cell* **12,** 85–100.

Grady, R. M., Grange, R. W., Lau, K. S., Maimone, M. M., Nichol, M. C., Stull, J. T., and Sanes, J. R. (1999). Role for alpha-dystrobrevin in the pathogenesis of dystrophin-dependent muscular dystrophies. *Nat. Cell Biol.* **1,** 215–220.

Granger, B. L., and Lazarides, E. (1980). Synemin: A new high molecular weight protein associated with desmin and vimentin filaments in muscle. *Cell* **22,** 727–738.

Granger, B. L., and Lazarides, E. (1982). Structural association of synemin and vimentin in avian erythrocytes revealed by immunoelectron microscopy. *Cell* **30,** 263–275.

Grant, P., Sharma, P., and Pant, H. C. (2001). Cyclin-dependent protein kinase 5 (Cdk5) and the regulation of neurofilament metabolism. *Eur. J. Biochem.* **268,** 1534–1546.

Green, K. J., Parry, D. A. D., Steinert, P. M., Virata, M. L. A., Wagner, R. M., Angst, B. D., and Nilles, L. A. (1990). Structure of the human desmoplakins: Implications for function in the desmosomal plaque. *J. Biol. Chem.* **265,** 2603–2612.

Green, K. J., Stappenbeck, T. S., Parry, D. A. D., and Virata, M. L. A. (1992a). Structure of desmoplakin and its association with intermediate filaments. *J. Dermatol.* **19,** 765–769.

Green, K. J., Virata, M. L. A., Elgart, G. W., Stanley, J. R., and Parry, D. A. D. (1992b). Comparative structural analysis of desmoplakin, bullous pemphigoid antigen and plectin: Members of a new gene family involved in organization of intermediate filaments. *Int. J. Biol. Macromol.* **14,** 145–153.

Groot, K. R., Sevilla, L. M., Nishi, K., DiColandrea, T., and Watt, F. M. (2004). Kazrin, a novel periplakin-interacting protein associated with desmosomes and the keratinocyte plasma membrane. *J. Cell Biol.* **166,** 653–659.

Guo, L., Degenstein, L., Dowling, J., Yu, Q.-C., Wollmann, R., Perman, B., and Fuchs, E. (1995). Gene targeting of BPAG1: Abnormalities in mechanical strength and cell migration in stratified epithelia and neurologic degeneration. *Cell* **81,** 233–243.

Harding, H. W., and Rogers, G. E. (1971). Epsilon-(gamma-glutamyl)lysine cross-linkage in citrulline-containing protein fractions from hair. *Biochemistry* **10,** 624–630.

Harding, H. W., and Rogers, G. E. (1972). Formation of the -(-glutamyl) lysine cross-link in hair proteins. Investigation of transamidases in hair follicles. *Biochemistry* **11,** 2858–2863.

Harding, H. W., and Rogers, G. E. (1976). Isolation of peptides containing citrulline and the cross-link, epsilon-(gamma-glutamyl)lysine, from hair medulla protein. *Biochim. Biophys. Acta* **427,** 315–324.

Hatzfeld, M. (1999). The armadillo family of structural proteins. *Int. Rev. Cytol.* **186,** 179–224.

Hatzfeld, M., Green, K. J., and Sauter, H. (2003). Targeting of p0071 to desmosomes and adherens junctions is mediated by different protein domains. *J. Cell Sci.* **116,** 1219–1233.

Hatzfeld, M., Haffner, C., Schulze, K., and Vinzens, U. (2000). The function of plakophilin 1 in desmosome assembly and actin filament organization. *J. Cell Biol.* **149,** 209–222.

Hatzfeld, M., Kristjansson, G. I., Plessmann, U., and Weber, K. (1994). Band 6 protein, a major constituent of desmosomes from stratified epithelia, is a novel member of the armadillo multigene family. *J. Cell Sci.* **107,** 2259–2270.

Helfand, B. T., Chang, L., and Goldman, R. D. (2004). Intermediate filaments are dynamic and motile elements of cellular architecture. *J. Cell Sci.* **117,** 133–141.

Helfand, B. T., Loomis, P., Yoon, M., and Goldman, R. D. (2003). Rapid transport of neural intermediate filament protein. *J. Cell Sci.* **116,** 2345–2359.

Helfand, B. T., Mikami, A., Vallee, R. B., and Goldman, R. D. (2002). A requirement for cytoplasmic dynein and dynactin in intermediate filament network assembly and organization. *J. Cell Biol.* **157,** 795–806.

Hemken, P. M., Bellin, R. M., Sernett, S. W., Becker, B., Huiatt, T. W., and Robson, R. M. (1997). Molecular characteristics of the novel intermediate filament protein paranemin. Sequence reveals EAP-300 and IFAPa-400 are highly homologous to paranemin. *J. Biol. Chem.* **272,** 32489–32499.

Herrmann, H., and Aebi, U. (2000). Intermediate filaments and their associates: Multi-talented structural elements specifying cytoarchitecture and cytodynamics. *Curr. Opin. Cell Biol.* **12,** 79–90.

Herrmann, H., and Foisner, R. (2003). Intermediate filaments: Novel assembly models and exciting new functions for nuclear lamins. *Cell Mol. Life Sci.* **60,** 1607–1612.

Herrmann, H., and Wiche, G. (1987). Plectin and IFAP-300K are homologous proteins binding to microtubule-associated proteins 1 and 2 and to the 240 kilodalton subunit of spectrin. *J. Biol. Chem.* **262,** 1320–1325.

Hesse, M., Magin, T. M., and Weber, K. (2001). Genes for intermediate filament proteins and the draft sequence of the human genome: Novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *J. Cell Sci.* **114,** 2569–2575.

Hijikata, T., Murakami, T., Imamura, M., Fujimaki, N., and Ishikawa, H. (1999). Plectin is a linker of intermediate filaments to Z-discs in skeletal muscle fibers. *J. Cell Sci.* **112,** 867–876.

Hijikata, T., Murakami, T., Ishikawa, H., and Yorifuji, H. (2003). Plectin tethers desmin intermediate filaments onto subsarcolemmal dense plaques containing dystrophin and vinculin. *Histochem. Cell Biol.* **119,** 109–123.

Hofmann, I., Mertens, C., Brettel, M., Nimmrich, V., Schnolzer, N., and Herrmann, H. (2000). Interaction of plakophilins with desmoplakin and intermediate filament proteins: An *in vitro* analysis. *J. Cell Sci.* **113,** 2471–2483.

Holzbaur, E. L. (2004). Tangled NUDELs? *Nat. Cell Biol.* **6,** 569–570.

Hopkinson, S. B., and Jones, J. C. (2000). The N terminus of the transmembrane protein BP180 interacts with the N-terminal domain of BP230, thereby mediating keratin cytoskeleton anchorage to the cell surface at the site of the hemidesmosome. *Mol. Biol. Cell* **11,** 277–286.

Houseweart, M. K., and Cleveland, D. W. (1999). Cytoskeletal linnkers: New MAPs for old destinations. *Curr. Biol.* **9,** R864–866.

Howman, E. V., Sullivan, N., Poon, E. P., Britton, J. E., Hilton-Jones, D., and Davies, K. E. (2003). Syncoilin accumulation in two patients with desmin-related myopathy. *Neuromuscul. Disord.* **13,** 42–48.

Huen, A. C., Park, J. K., Godsel, L. M., Chen, X., Bannon, L. J., Amargo, E. V., Hudson, T. Y., Mongiu, A. K., Leigh, I. M., Kelsell, D. P., Gumbiner, B. M., and Green, K. J. (2002). Intermediate filament-membrane attachments function synergistically with actin-dependent contacts to regulate intercellular adhesive strength. *J. Cell Biol.* **159,** 1005–1017.

Hynes, R. O. (2002). Integrins: Bidirectional, allosteric signaling machines. *Cell* **110,** 673–687.

Inada, H., Izawa, I., Nishizawa, M., Fujita, E., Kiyono, T., Takahashi, T., Momoi, T., and Inagaki, M. (2001). Keratin attenuates tumor necrosis factor-induced cytotoxicity through association with TRADD. *J. Cell Biol.* **155,** 415–426.

Ishida-Yamamoto, A., Hashimoto, Y., Manabe, M., O'Guin, W. M., Dale, B. A., and Iizuka, H. (1997). Distinctive expression of filaggrin and trichohyalin during various pathways of epithelial differentiation. *Br. J. Dermatol.* **137,** 9–16.

Izawa, I., Nishizawa, M., Ohtakara, K., Ohtsuka, K., Inada, H., and Inagaki, M. (2000). Identification of Mrj, a DnaJ/Hsp40 family protein, as a keratin 8/18 filament regulatory protein. *J. Biol. Chem.* **275,** 34521–34527.

Jefferson, J. J., Leung, C. L., and Liem, R. K. (2004). Plakins: Goliaths that link cell junctions and the cytoskeleton. *Nat. Rev. Mol. Cell. Biol.* **5,** 542–553.

Jones, J. C., Hopkinson, S. B., and Goldfinger, L. E. (1998). Structure and assembly of hemidesmosomes. *Bioessays* **20,** 488–494.

Jonkman, M. F. (1999). Hereditary skin diseases of hemidesmosomes. *J. Dermatol. Sci.* **20,** 103–121.

Kalinin, A. E., Idler, W. W., Marekov, L. N., McPhie, P., Bowers, B., Steinert, P. M., and Steven, A. C. (2004). Co-assembly of envoplakin and periplakin into oligomers and Ca(2+)-dependent vesicle binding: Implications for cornified cell envelope formation in stratified squamous epithelia. *J. Biol. Chem.* **279,** 22773–22780.

Kapprell, H.-P., Owaribe, K., and Franke, W. W. (1988). Identification of a basic protein of Mr 75,000 as an accessory desmosomal plaque protein in stratified and complex epithelia. *J. Biol. Chem.* **106,** 1679–1691.

Karashima, T., and Watt, F. M. (2002). Interaction of periplakin and envoplakin with intermediate filaments. *J. Cell Sci.* **115,** 5027–5037.

Kazerounian, S., Uitto, J., and Aho, S. (2002). Unique role for the periplakin tail in intermediate filament association: Specific binding to keratin 8 and vimentin. *Exp. Dermatol.* **11,** 428–438.

Kazmierski, S. T., Antin, P. B., Witt, C. C., Huebner, N., McElhinny, A. S., Labeit, S., and Gregorio, C. C. (2003). The complete mouse nebulin gene sequence and the identification of cardiac nebulin. *J. Mol. Biol.* **328,** 835–846.

Kim, E., Arnould, T., Sellin, L., Benzing, T., Comella, N., Kocher, O., Tsiokas, L., Sukhatme, V. P., and Walz, G. (1999a). Interaction between RGS7 and polycystin. *Proc. Natl. Acad. Sci. USA* **96,** 6371–6376.

Kim, E., Arnould, T., Sellin, L. K., Benzing, T., Fan, M. J., Gruning, W., Sokol, S. Y., Drummond, I., and Walz, G. (1999b). The polycystic kidney disease 1 gene product modulates Wnt signaling. *J. Biol. Chem.* **274,** 4947–4953.

Kiyokawa, C., Ruhrberg, C., Nie, Z., Karashima, T., Mori, O., Nishikawa, T., Green, K. J., Anhalt, G. J., DiColandrea, T., Watt, F. M., and Hashimoto, T. (1998). Envoplakin and periplakin are components of the paraneoplastic pemphigus antigen complex. *J. Invest. Derm.* **111,** 1236–1238.

Koster, J., Borradori, L., and Sonnenberg, A., (2004a). Hemidesmosomes: Molecular organization and their importance for cell adhesion and disease. *In* ''Handbook of Experimental Pharmacology'' (K. Starke, Ed.), Vol. 165, pp.243–280. Springer-Verlag, Berlin.

Koster, J., Geerts, D., Favre, B., Borradori, L., and Sonnenberg, A. (2003). Analysis of the interactions between BP180, BP230, plectin, and the integrin alpha6beta4 important for hemidesmosome assembly. *J. Cell Sci.* **116,** 387–399.

Koster, J., van Wilpe, S., Kuikman, I., Litjens, S. H., and Sonnenberg, A. (2004b). Role of binding of plectin to the integrin beta4 subunit in the assembly of hemidesmosomes. *Mol. Biol. Cell* **15,** 1211–1223.

Kouklis, P. D., Hutton, E., and Fuchs, E. (1994). Making a connection: Direct binding between keratin intermediate filaments and desmosomal proteins. *J. Biol. Chem.* **127,** 1049–1060.

Kowalczyk, A. P., Bornslaeger, E. A., Borgwardt, J. E., Palka, H. L., Dhaliwal, A. S., Corcoran, C. M., Denning, M. F., and Green, K. J. (1997). The amino-terminal domain of desmoplakin binds to plakoglobin and clusters desmosomal cadherin-plakoglobin complexes. *J. Biol. Chem.* **139,** 773–784.

Kowalczyk, A. P., Hatzfeld, M., Bornslaeger, E. A., Kopp, D. S., Borgwardt, J. E., Corcoran, C. M., Settler, A., and Green, K. J. (1999). The head domain of plakophilin-1 binds to and enhances its recruitment to desmosomes: Implications for cutaneous disease. *J. Biol. Chem.* **274,** 18145–18148.

Kruger, M., Wright, J., and Wang, K. (1991). Nebulin as a length regulator of thin filaments of vertebrate skeletal muscles: Correlation of thin filament length, nebulin size, and epitope profile. *J. Cell Biol.* **115,** 97–107.

Ku, N. O., Azhar, S., and Omary, M. B. (2002a). Keratin 8 phosphorylation by p38 kinase regulates cellular keratin filament reorganization: Modulation by a keratin 1-like disease causing mutation. *J. Biol. Chem.* **277,** 10775–10782.

Ku, N. O., Darling, J. M., Krams, S. M., Esquivel, C. O., Keeffe, E. B., Sibley, R. K., Lee, Y. M., Wright, T. L., and Omary, M. B. (2003a). Keratin 8 and 18 mutations are risk factors for developing liver disease of multiple etiologies. *Proc. Natl. Acad. Sci. USA* **100,** 6063–6068.

Ku, N. O., Fu, H., and Omary, M. B. (2004). Raf-1 activation disrupts its binding to keratins during cell stress. *J. Cell Biol.* **166,** 479–485.

Ku, N. O., Gish, R., Wright, T. L., and Omary, M. B. (2001). Keratin 8 mutations in patients with cryptogenic liver disease. *N. Engl. J. Med.* **344,** 1580–1587.

Ku, N. O., Liao, J., and Omary, M. B. (1998). Phosphorylation of human keratin 18 serine 33 regulates binding to 14-3-3 proteins. *EMBO J.* **17,** 1892–1906.

Ku, N. O., Michie, S., Resurreccion, E. Z., Broome, R. L., and Omary, M. B. (2002b). Keratin binding to 14-3-3 proteins modulates keratin filaments and hepatocyte mitotic progression. *Proc. Natl. Acad. Sci. USA* **99,** 4373–4378.

Ku, N. O., Michie, S. A., Soetikno, R. M., Resurreccion, E. Z., Broome, R. L., Oshima, R. G., and Omary, M. B. (1996). Susceptibility to hepatotoxicity in transgenic mice that express a dominant-negative human keratin 18 mutant. *J. Clin. Invest.* **98,** 1034–1046.

Ku, N. O., Soetikno, R. M., and Omary, M. B. (2003b). Keratin mutation in transgenic mice predisposes to Fas but not TNF-induced apoptosis and massive liver injury. *Hepatology* **37,** 1006–1014.

Kuechle, M. K., Thulin, C. D., Presland, R. B., and Dale, B. A. (1999). Profilaggrin requires both linker and filaggrin peptide sequences to form granules: Implications for profilaggrin processing *in vivo*. *J. Invest. Dermatol.* **112,** 843–852.

Kuhn, F., Lassing, C., Range, A., Mueller, M., Hunziker, T., Ziemiecki, A., and Andres, A. C. (1999). Pmg-1 and pmg-2 constitute a novel family of KAP genes differentially expressed during skin and mammary gland development. *Mech. Dev.* **86,** 193–196.

Langley, R. C., Jr., and Cohen, C. M. (1986). Association of spectrin with desmin intermediate filaments. *J. Cell Biochem.* **30,** 101–109.

Langley, R. C., Jr., and Cohen, C. M. (1987). Cell type-specific association between two types of spectrin and two types of intermediate filaments. *Cell Motil. Cytoskel.* **8,** 165–173.

Law, S. F., O'Neill, G. M., Fashena, S. J., Einarson, M. B., and Golemis, E. A. (2000). The docking protein HEF1 is an apoptotic mediator at focal adhesion sites. *Mol. Cell. Biol.* **20,** 5184–5195.

Lazebnik, Y. A., Takahashi, A., Moir, R. D., Goldman, R. D., Poirier, G. G., Kaufmann, S. H., and Earnshaw, W. C. (1995). Studies of the lamin proteinase reveal multiple parallel biochemical pathways during apoptotic execution. *Proc. Natl. Acad. Sci. USA* **92,** 9042–9046.

Lee, J. C., Schickling, O., Stegh, A. H., Oshima, R. G., Dinsdale, D., Cohen, G. M., and Peter, M. E. (2002). DEDD regulates degradation of intermediate filaments during apoptosis. *J. Cell Biol.* **158,** 1051–1066.

Lee, M. K., and Cleveland, D. W. (1996). Neuronal intermediate filaments. *Annu. Rev. Neurosci.* **19,** 187–217.

Lee, S.-C., Kim, I.-G., Marekov, L. N., O'Keefe, E. J., Parry, D. A. D., and Steinert, P. M. (1993). The structure of human trichohyalin: Potential multiple roles as a functional EF-hand-like calcium-binding protein, a cornified cell envelope precursor, and an intermediate filament-associated (cross-linking) protein. *J. Biol. Chem.* **268,** 12164–12176.

Leung, C. L., Green, K. J., and Liem, R. K. (2002). Plakins: A family of versatile cytolinker proteins. *Trends Cell Biol.* **12,** 37–45.

Leung, C. L., Liem, R. K., Parry, D. A., and Green, K. J. (2001a). The plakin family. *J. Cell Sci.* **114,** 3409–3410.

Leung, C. L., Sun, D., and Liem, R. K. (1999). The intermediate filament protein peripherin is the specific interaction partner of mouse BPAG1-n (dystonin) in neurons. *J. Cell Biol.* **144,** 435–446.

Leung, C. L., Zheng, M., Prater, S. M., and Liem, R. K. (2001b). The BPAG1 locus: Alternative splicing produces multiple isoforms with distinct cytoskeletal linker domains, including predominant isoforms in neurons and muscles. *J. Cell Biol.* **154,** 691–698.

Liao, G., and Gundersen, G. G. (1998). Kinesin is a candidate for cross-bridging microtubules and intermediate filaments. Selective binding of kinesin to detyrosinated tubulin and vimentin. *J. Biol. Chem.* **273,** 9797–9803.

Liao, J., Lowthert, L. A., Ghori, N., and Omary, M. B. (1995). The 70-kDa heat shock proteins associate with glandular intermediate filaments in an ATP-dependent manner. *J. Biol. Chem.* **270,** 915–922.

Liao, J., and Omary, M. B. (1996). 14-3-3 proteins associate with phosphorylated simple epithelial keratins during cell cycle progression and act as a solubility cofactor. *J. Cell. Biol.* **133,** 345–357.

Liao, J., Price, D., and Omary, M. B. (1997). Association of glucose-regulated protein (grp78) with human keratin 8. *FEBS Lett.* **417,** 316–320.

Liebowitz, D., Kopan, R., Fuchs, E., Sample, J., and Kieff, E. (1987). An Epstein-Barr virus transforming protein associates with vimentin in lymphocytes. *Mol. Cell. Biol.* **7,** 2299–2308.

Lieska, N., Yang, H.-Y., and Goldman, R. D. (1985). Purification of the 300K intermediate filament-associated protein and its in vitro recombination with intermediate filaments. *J. Biol. Chem.* **101,** 802–813.

Lin, C. S., Park, T., Chen, Z. P., and Leavitt, J. (1993). Human plastin genes. Comparative gene structure, chromosome location, and differential expression in normal and neoplastic cells. *J. Biol. Chem.* **268,** 2781–2792.

Lin, C. S., Shen, W., Chen, Z. P., Tu, Y. H., and Matsudaira, P. (1994). Identification of I-plastin, a human fimbrin isoform expressed in intestine and kidney. *Mol. Cell. Biol.* **14,** 2457–2467.

Liu, C. G., Maercker, C., Castanon, M. J., Hauptmann, R., and Wiche, G. (1996). Human plectin: Organization of the gene, sequence analysis, and chromosome localization (8q24). *Proc. Natl. Acad. Sci.* **93,** 4278–4283.

Lonsdale-Eccles, J. D., Teller, D. C., and Dale, B. A. (1982). Characterization of a phosphorylated form of the intermediate filament-aggregating protein filaggrin. *Biochemistry* **21,** 5940–5948.

Maatta, A., DiColandrea, T., Groot, K., and Watt, F. M. (2001). Gene targeting of envoplakin, a cytoskeletal linker protein and precursor of the epidermal cornified envelope. *Mol. Cell. Biol.* **21,** 7047–7053.

Mabuchi, K., Li, B., Ip, W., and Tao, T. (1997). Association of calponin with desmin intermediate filaments. *J. Biol. Chem.* **272,** 22662–22666.

Mabuchi, K., Li, Y., Tao, T., and Wang, C. L. (1996). Immunocytochemical localization of caldesmon and calponin in chicken gizzard smooth muscle. *J. Muscle Res. Cell Motil.* **17,** 243–260.

Mack, J. W., Steven, A. C., and Steinert, P. M. (1993). The mechanism of interaction of filaggrin with intermediate filaments. The ionic zipper hypothesis. *J. Mol. Biol.* **232,** 50–66.

Mahoney, M. G., Aho, S., Uitto, J., and Stanley, J. R. (1998). The members of the plakin family of proteins recognized by paraneoplastic pemphigus antibodies include periplakin. *J. Invest. Derm.* **111,** 308–313.

Manabe, M., and O'Guin, W. M. (1994). Existence of trichohyalin-keratohyalin hybrid granules: Co-localization of two major intermediate filament-associated proteins in non-follicular epithelia. *Differentiation* **58,** 65–75.

Manabe, M., Sanchez, M., Sun, T. T., and Dale, B. A. (1991). Interaction of filaggrin with keratin filaments during advanced stages of normal human epidermal differentiation and in ichthyosis vulgaris. *Differentiation* **48,** 43–50.

Mangeat, P. H., and Burridge, K. (1984). Immunoprecipitation of nonerythrocyte spectrin within live cells following microinjection of specific antibodies: Relation to cytoskeletal structures. *J. Cell. Biol.* **98,** 1363–1377.

Marceau, N., Loranger, A., Gilbert, S., Daigle, N., and Champetier, S. (2001). Keratin-mediated resistance to stress and apoptosis in simple epithelial cells in relation to health and disease. *Biochem. Cell Biol.* **79,** 543–555.

Marekov, L. N., and Steinert, P. M. (1998). Ceramides are bound to structural proteins of the human foreskin epidermal cornified cell envelope. *J. Biol. Chem.* **273,** 17763–17770.

Markova, N. G., Marekov, L. N., Chipev, C. C., Gan, S. Q., Idler, W. W., and Steinert, P. M. (1993). Profilaggrin is a major epidermal calcium-binding protein. *Mol. Cell. Biol.* **13,** 613–625.

McElhinny, A. S., Kazmierski, S. T., Labeit, S., and Gregorio, C. C. (2003). Nebulin: The nebulous, multifunctional giant of striated muscle. *Trends Cardiovasc. Med.* **13,** 195–201.

McElhinny, A. S., Kolmerer, B., Fowler, V. M., Labeit, S., and Gregorio, C. C. (2001). The N-terminal end of nebulin interacts with tropomodulin at the pointed ends of the thin filaments. *J. Biol. Chem.* **276,** 583–592.

McLean, W. H. I., Pulkkinen, L., Smith, F. J. D., Rugg, E. L., Lane, E. B., Bullrich, F., Burgeson, R. E., Amano, S., Hudson, D. L., Owaribe, K., McGrath, J.A., McMillan, J.R., Eady, R.A.J., Leigh, I.M., Christiano, A.M., and Uitto, J. (1996). Loss of plectin causes epidermolysis bullosa with muscular dystrophy: cDNA cloning and genomic organization. *Genes Dev.* **10,** 1724–1735.

McMillan, J. R., and Shimizu, H. (2001). Desmosomes: Structure and function in normal and diseased epidermis. *J. Dermatol.* **28,** 291–298.

Meng, J.-J., Bornslaeger, E. A., Green, K. J., Steinert, P. M., and Ip, W. (1997). Two hybrid analysis reveals fundamental differences in direct interactions between desmoplakin and cell type specific intermediate filaments. *J. Biol. Chem.* **272,** 21495–21503.

Miravet, S., Piedra, J., Castano, J., Raurell, I., Franci, C., Dunach, M., and Garcia de Herreros, A. (2003). Tyrosine phosphorylation of plakoglobin causes contrary effects on its association with desmosomes and adherens junction components and modulates beta-catenin-mediated transcription. *Mol. Cell. Biol.* **23,** 7391–7402.

Mizuno, Y., Guyon, J. R., Watkins, S. C., Mizushima, K., Sasaoka, T., Imamura, M., Kunkel, L. M., and Okamoto, K. (2004). beta-synemin localizes to regions of high stress in human skeletal myofibers. *Muscle Nerve.* **30,** 337–346.

Mizuno, Y., Thompson, T. G., Guyon, J. R., Lidov, H. G., Brosius, M., Imamura, M., Ozawa, E., Watkins, S. C., and Kunkel, L. M. (2001). Desmuslin, an intermediate filament protein that interacts with alpha-dystrobrevin and desmin. *Proc. Natl. Acad. Sci. USA* **98,** 6156–6161.

Murata, T., Goshima, F., Nishizawa, Y., Daikoku, T., Takakuwa, H., Ohtsuka, K., Yoshikawa, T., and Nishiyama, Y. (2002). Phosphorylation of cytokeratin 17 by herpes simplex virus type 2 US3 protein kinase. *Microbiol. Immunol.* **46,** 707–719.

Nakagawa, H., Ishihara, M., and Ohashi, K. (1993). 33-kDa peptides prepared from chicken gizzard smooth muscle bundle both actin and desmin filaments *in vitro. J. Biochem. (Tokyo)* **114,** 623–626.

Namba, Y., Ito, M., Zu, Y., Shigesada, K., and Maruyama, K. (1992). Human T cell L-plastin bundles actin filaments in a calcium-dependent manner. *J. Biochem. (Tokyo)* **112,** 503–507.

Napolitano, E. W., Pachter, J. S., Chin, S. S., and Liem, R. K. (1985). beta-Internexin, a ubiquitous intermediate filament-associated protein. *J. Cell Biol.* **101,** 1323–1331.

Nelson, W. J., and Nusse, R. (2004). Convergence of Wnt, beta-catenin, and cadherin pathways. *Science* **303,** 1483–1487.

Newey, S. E., Howman, E. V., Ponting, C. P., Benson, M. A., Nawrotzki, R., Loh, N. Y., Davies, K. E., and Blake, D. J. (2001). Syncoilin, a novel member of the intermediate filament superfamily that interacts with alpha-dystrobrevin in skeletal muscle. *J. Biol. Chem.* **276,** 6645–6655.

Nguyen, M. D., Shu, T., Sanada, K., Lariviere, R. C., Tseng, H. C., Park, S. K., Julien, J. P., and Tsai, L. H. (2004). A NUDEL-dependent mechanism of neurofilament assembly regulates the integrity of CNS neurons. *Nat. Cell Biol.* **6,** 595–608.

Nicholl, I. D., and Quinlan, R. A. (1994). Chaperone activity of alpha-crystallins modulates intermediate filament assembly. *EMBO J.* **13,** 945–953.

Nievers, M. G., Schaapveld, R. Q., and Sonnenberg, A. (1999). Biology and function of hemidesmosomes. *Matrix Biol.* **18,** 5–17.

Nikolic, B., MacNulty, E., Mir, B., and Wiche, G. (1996). Basic amino acid residue cluster within nuclear targeting sequence motif is essential for cytoplasmic plectin-vimentin network junctions. *J. Cell Biol.* **134,** 1455–1467.

Nishida, W., Abe, M., Takahashi, K., and Hiwada, K. (1990). Do thin filaments of smooth muscle contain calponin? A new method for the preparation. *FEBS Lett.* **268,** 165–168.

Norgett, E. E., Hatsell, S. J., Carvajal-Huerta, L., Cabezas, J. C., Common, J., Purkis, P. E., Whittock, N., Leigh, I. M., Stevens, H. P., and Kelsell, D. P. (2000). Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair, and keratoderma. *Hum. Mol. Genet.* **9,** 2761–2766.

North, A. J., Bardsley, W. G., Hyam, J., Bornslaeger, E. A., Cordingley, H. C., Trinnaman, B., Hatzfeld, M., Green, K. J., Magee, A. I., and Garrod, D. R. (1999). Molecular map of the desmosomal plaque. *J. Cell Sci.* **112,** 4325–4336.

North, A. J., Gimona, M., Cross, R. A., and Small, J. V. (1994). Calponin is localised in both the contractile apparatus and the cytoskeleton of smooth muscle cells. *J. Cell Sci.* **107,** 437–444.

O'Guin, W. M., Sun, T. T., and Manabe, M. (1992). Interaction of trichohyalin with intermediate filaments: Three immunologically defined stages of trichohyalin maturation. *J. Invest. Dermatol.* **98,** 24–32.

O'Keefe, E. J., Erickson, H. P., and Bennett, V. (1989). Desmoplakin I and desmoplakin II: Purification and characterization. *J. Biol. Chem.* **264,** 8310–8318.

O'Keefe, E. J., Hamilton, E. H., Lee, S. C., and Steinert, P. (1993). Trichohyalin: A structural protein of hair, tongue, nail, and epidermis. *J. Invest. Dermatol.* **101,** 65S–71S.

Okumura, M., Uematsu, J., Hirako, Y., Nishizawa, Y., Shimizu, H., Kido, N., and Owaribe, K. (1999). Identification of the hemidesmosomal 500 kDa protein (HD1) as plectin. *J. Biochem. (Tokyo)* **126,** 1144–1150.

Okumura, M., Yamakawa, H., Ohara, O., and Owaribe, K. (2002). Novel alternative splicings of BPAG1 (bullous pemphigoid antigen 1) including the domain structure closely related to MACF (microtubule actin cross-linking factor). *J. Biol. Chem.* **277,** 6682–6687.

O'Neill, G. M., Fashena, S. J., and Golemis, E. A. (2000). Integrin signalling: A new Cas(t) of characters enters the stage. *Trends Cell Biol.* **10,** 111–119.

Oshima, R. G. (2002). Apoptosis and keratin intermediate filaments. *Cell Death Differ.* **9,** 486–492.

Osmanagic-Myers, S., and Wiche, G. (2004). Plectin-RACK1 (receptor for activated C kinase 1) scaffolding: A novel mechanism to regulate protein kinase C activity. *J. Biol. Chem.* **279,** 18701–18710.

Ouyang, P., and Sugrue, S. P. (1996). Characterization of pinin, a novel protein associated with the desmosome-intermediate filament complex. *J. Biol. Chem.* **135,** 1027–1042.

Ouyang, P., Zhen, Y. Y., and Sugrue, S. P. (1997). Cloning and analysis of cDNA encoding murine pinin. *Gene* **197,** 115–120.

Owens, D. W., Wilson, N. J., Hill, A. J., Rugg, E. L., Porter, R. M., Hutcheson, A. M., Quinlan, R. A., van Heel, D., Parkes, M., Jewell, D. P., Campbell, S. S., Ghosh, S., Satsangi, J., and Lane, E. B. (2004). Human keratin 8 mutations that disturb filament assembly observed in inflammatory bowel disease patients. *J. Cell Sci.* **117,** 1989–1999.

Palka, H. L., and Green, K. J. (1997). Roles of plakoglobin end domains in desmosome assembly. *J. Cell Sci.* **110,** 2359–2371.

Paramio, J. M., Casanova, M. L., Segrelles, C., Mittnacht, S., Lane, E. B., and Jorcano, J. L. (1999). Modulation of cell proliferation by cytokeratins K10 and K16. *Mol. Cell. Biol.* **19,** 3086–3094.

Paramio, J. M., Segrelles, C., Ruiz, S., and Jorcano, J. L. (2001). Inhibition of protein kinase B (PKB) and PKCzeta mediates keratin K10-induced cell cycle arrest. *Mol. Cell. Biol.* **21,** 7449–7459.

Pardo, J. V., Siliciano, J. D., and Craig, S. W. (1983). A vinculin-containing cortical lattice in skeletal muscle: Transverse lattice elements (''costameres'') mark sites of attachment between myofibrils and sarcolemma. *Proc. Natl. Acad. Sci.* **80,** 1081–1012.

Pearton, D. J., Nirunsuksiri, W., Rehemtulla, A., Lewis, S. P., Presland, R. B., and Dale, B. A. (2001). Proprotein convertase expression and localization in epidermis: Evidence for multiple roles and substrates. *Exp. Dermatol.* **10,** 193–203.

Peifer, M., McCrea, P. D., Green, K. J., Wieschaus, E., and Gumbiner, B. M. (1992). The vertebrate adhesive junction proteins b-catenin and plakoglobin and the

Drosophila segment polarity gene armadillo form a multigene family with similar properties. *J. Biol. Chem.* **118,** 681–691.

Pelin, K., Hilpela, P., Donner, K., Sewry, C., Akkari, P. A., Wilton, S. D., Wattanasirichaigoon, D., Bang, M. L., Centner, T., Hanefeld, F., Odent, S., Fardeau, M., Urtizberea, J. A., Muntoni, F., Dubowitz, V., Beggs, A. H., Laing, N. G., Labeit, S., de la Chapelle, A., and Wallgren-Pettersson, C. (1999). Mutations in the nebulin gene associated with autosomal recessive nemaline myopathy. *Proc. Natl. Acad. Sci. USA* **96,** 2305–2310.

Perng, M. D., Cairns, L., van den, I. P., Prescott, A., Hutcheson, A. M., and Quinlan, R. A. (1999a). Intermediate filament interactions can be altered by HSP27 and alphaB-crystallin. *J. Cell Sci.* **112,** 2099–2112.

Perng, M. D., Muchowski, P. J., van Den, I. P., Wu, G. J., Hutcheson, A. M., Clark, J. I., and Quinlan, R. A. (1999b). The cardiomyopathy and lens cataract mutation in alphaB-crystallin alters its protein structure, chaperone activity, and interaction with intermediate filaments *in vitro. J. Biol. Chem.* **274,** 33235–33243.

Poon, E., Howman, E. V., Newey, S. E., and Davies, K. E. (2002). Association of syncoilin and desmin: Linking intermediate filament proteins to the dystrophin-associated protein complex. *J. Biol. Chem.* **277,** 3433–3439.

Powell, B. C., and Rogers, G. E. (1997). The role of keratin proteins and their genes in the growth, structure and properties of hair. *In* ''Formation and structure of human hair'' (P. Jollès, H. Zahn, and H. Höcker, Eds.), pp. 59–148. Basel; Birkhäuser Verlag, Boston.

Prahlad, V., Helfand, B. T., Langford, G. M., Vale, R. D., and Goldman, R. D. (2000). Fast transport of neurofilament protein along microtubules in squid axoplasm. *J. Cell Sci.* **113,** 3939–3946.

Presland, R. B., Bassuk, J. A., Kimball, J. R., and Dale, B. A. (1995). Characterization of two distinct calcium-binding sites in the amino-terminus of human profilaggrin. *J. Invest. Dermatol.* **104,** 218–223.

Presland, R. B., and Dale, B. A. (2000). Epithelial structural proteins of the skin and oral cavity: Function in health and disease. *Crit. Rev. Oral Biol. Med.* **11,** 383–408.

Presland, R. B., Haydock, P. V., Fleckman, P., Nirunsuksiri, W., and Dale, B. A. (1992). Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. *J. Biol. Chem.* **267,** 23772–23781.

Presland, R. B., Kuechle, M. K., Lewis, S. P., Fleckman, P., and Dale, B. A. (2001). Regulated expression of human filaggrin in keratinocytes results in cytoskeletal disruption, loss of cell-cell adhesion, and cell cycle arrest. *Exp. Cell Res.* **270,** 199–213.

Price, M., and Sanger, J. W. (1979). Intermediate filaments connect Z-discs in adult chicken muscle. *J. Exp. Zool.* **208,** 263–269.

Pulkkinen, L., Smith, F. J. D., Shimizu, H., Murata, S., Yaoita, H., Hachisuka, H., Nishikawa, T., McLean, W. H. I., and Uitto, J. (1996). Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy. *Hum. Molec. Gent.* **5,** 1539–1546.

Pytela, R., and Wiche, G. (1980). High molecular weight polypeptides (270,000–340,000) from cultured cells are related to hog brain microtubule-associated proteins but copurify with intermediate filaments. *Proc. Natl. Acad. Sci. USA* **77,** 4808–4812.

Qi, Z., Tang, D., Zhu, X., Fujita, D. J., and Wang, J. H. (1998). Association of neurofilament proteins with neuronal Cdk5 activator. *J. Biol. Chem.* **273,** 2329–2335.

Rampazzo, A., Nava, A., Malacrida, S., Beffagna, G., Bauce, B., Rossi, V., Zimbello, R., Simionati, B., Basso, C., Thiene, G., Towbin, J. A., and Danieli, G. A. (2002). Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. *Am. J. Hum. Genet.* **71,** 1200–1206.

Rao, M. V., Engle, L. J., Mohan, P. S., Yuan, A., Qiu, D., Cataldo, A., Hassinger, L., Jacobsen, S., Lee, V. M., Andreadis, A., Julien, J. P., Bridgman, P. C., and Nixon, R. A. (2002). Myosin Va binding to neurofilaments is essential for correct myosin Va distribution and transport and neurofilament density. *J. Cell Biol.* **159,** 279–290.

Rawlings, A. V., and Harding, C. R. (2004). Moisturization and skin barrier function. *Dermatol. Ther.* **17,** S43–S48.

Resing, K. A., Thulin, C., Whiting, K., al-Alawi, N., and Mostad, S. (1995). Characterization of profilaggrin endoproteinase 1. A regulated cytoplasmic endoproteinase of epidermis. *J. Biol. Chem.* **270,** 28193–28198.

Rezniczek, G. A., Abrahamsberg, C., Fuchs, P., Spazierer, D., and Wiche, G. (2003). Plectin 5′-transcript diversity: Short alternative sequences determine stability of gene products, initiation of translation and subcellular localization of isoforms. *Hum. Mol. Genet.* **12,** 3181–3194.

Rezniczek, G. A., Pereda, J. M. d., Reipert, S., and Wiche, G. (1998). Linking integrin $\alpha 6\beta 4$-based cell adhesion to the intermediate filament cytoskeleton: Direct interaction between the $\beta 4$ subunit and plectin at multiple molecular sites. *J. Biol. Chem.* **141,** 209–225.

Rogers, G. E., and Powell, B. C. (1993). Organization and expression of hair follicle genes. *J. Invest. Dermatol.* **101,** 50S–55S.

Rogers, M. A., Langbein, L., Winter, H., Beckmann, I., Praetzel, S., and Schweizer, J. (2004). Hair keratin associated proteins: Characterization of a second high sulfur KAP gene domain on human chromosome 21. *J. Invest. Dermatol.* **122,** 147–158.

Rogers, M. A., Langbein, L., Winter, H., Ehmann, C., Praetzel, S., and Schweizer, J. (2002). Characterization of a first domain of human high glycine-tyrosine and high sulfur keratin-associated protein (KAP) genes on chromosome 21q22.1. *J. Biol. Chem.* **277,** 48993–49002.

Roitbak, T., Ward, C. J., Harris, P. C., Bacallao, R., Ness, S. A., and Wandinger-Ness, A. (2004). A polycystin-1 multiprotein complex is disrupted in polycystic kidney disease cells. *Mol. Biol. Cell* **15,** 1334–1346.

Roper, K., and Brown, N. H. (2003). Maintaining epithelial integrity: A function for gigantic spectraplakin isoforms in adherens junctions. *J. Cell Biol.* **162,** 1305–1315.

Roper, K., Gregory, S. L., and Brown, N. H. (2002). The 'spectraplakins': Cytoskeletal giants with characteristics of both spectrin and plakin families. *J. Cell Sci.* **115,** 4215–4225.

Rothnagel, J. A., and Rogers, G. E. (1986). Trichohyalin, an intermediate filament-associated protein of the hair follicle. *J. Cell Biol.* **102,** 1419–1429.

Ruhrberg, C., Hajibagheri, M. A. N., Parry, D. A. D., and Watt, F. M. (1997). Periplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. *J. Biol. Chem.* **139,** 1835–1849.

Ruhrberg, C., Hajibagheri, M. A. N., Simon, M., Dooley, T. P., and Watt, F. M. (1996). Envoplakin, a novel precursor of the cornified envelope that has homology to desmoplakin. *J. Biol. Chem.* **134,** 715–729.

Ruhrberg, C., and Watt, F. M. (1997). The plakin family: Versatile organisers of cytoskeletal architecture. *Curr. Opin. Genet. Devel.* **7,** 392–397.

Sahlgren, C. M., Mikhailov, A., Vaittinen, S., Pallari, H. M., Kalimo, H., Pant, H. C., and Eriksson, J. E. (2003). Cdk5 regulates the organization of Nestin and its association with p35. *Mol. Cell. Biol.* **23,** 5090–5106.

Santos, M., Perez, P., Segrelles, C., Ruiz, S., Jorcano, J. L., and Paramio, J. M. (2003). Impaired NF-kappa B activation and increased production of tumor necrosis factor alpha in transgenic mice expressing keratin K10 in the basal layer of the epidermis. *J. Biol. Chem.* **278,** 13422–13430.

Sarria, A. J., Panini, S. R., and Evans, R. M. (1992). A functional role for vimentin intermediate filaments in the metabolism of lipoprotein-derived cholesterol in human SW-13 cells. *J. Biol. Chem.* **267,** 19455–19463.

Sawamura, D., Li, K., Chu, M.-L., and Uitto, J. (1991). Human bullous pemphigoid antigen (BPAG1). Amino acid sequences deduced from cloned cDNAs predict biologically important peptide segments and protein domains. *J. Biol. Chem.* **266,** 17784–17790.

Sawamura, D., Nomura, K., Sugita, Y., Mattei, M. G., Chu, M.L., Knowlton, R., and Uitto, J. (1990). Bullous pemphigoid antigen (BPAG1): cDNA cloning and mapping of the gene to the short arm of human chromosome 6. *Genomics* **8,** 722–726.

Scheffers, M. S., van der Bent, P., Prins, F., Spruit, L., Breuning, M. H., Litvinov, S. V., de Heer, E., and Peters, D. J. (2000). Polycystin-1, the product of the polycystic kidney disease 1 gene, co-localizes with desmosomes in MDCK cells. *Hum. Mol. Genet.* **9,** 2743–2750.

Schmelz, M., and Franke, W. W. (1993). Complexus adhaerentes, a new group of desmoplakin-containing junctions in endothelial cells: The syndesmos connecting retrothelial cells of lymph nodes. *Eur. J. Cell Biol.* **61,** 274–289.

Schröder, R., Furst, D. O., Klasen, C., Reimann, J., Herrmann, H., and van der Ven, P. F. (2000). Association of plectin with Z-discs is a prerequisite for the formation of the intermyofibrillar desmin cytoskeleton. *Lab. Invest.* **80,** 455–464.

Schwarz, M. A., Owaribe, K., Kartenbeck, J., and Franke, W. W. (1990). Desmosomes and hemidesmosomes: Constitutive molecular components. *Ann. Rev. Cell Biol.* **6,** 461–491.

Scott, C., Keating, L., Bellamy, M., and Baines, A. J. (2001). Protein 4.1 in forebrain postsynaptic density preparations: Enrichment of 4.1 gene products and detection of 4.1R binding proteins. *Eur. J. Biochem.* **268,** 1084–1094.

Scott, I. R., and Harding, C. R. (1986). Filaggrin breakdown to water binding compounds during development of the rat stratum corneum is controlled by the water activity of the environment. *Dev. Biol.* **115,** 84–92.

Seifert, G. J., Lawson, D., and Wiche, G. (1992). Immunolocalization of the intermediate filament-associated protein plectin at focal contacts and actin stress fibers. *Eur. J. Cell Biol.* **59,** 138–147.

Sevcik, J., Urbanikova, L., Kost'an, J., Janda, L., and Wiche, G. (2004). Actin-binding domain of mouse plectin. Crystal structure and binding to vimentin. *Eur. J. Biochem.* **271,** 1873–1884.

Shah, J. V., Flanagan, L. A., Janmey, P. A., and Leterrier, J. F. (2000). Bidirectional translocation of neurofilaments along microtubules mediated in part by dynein/dynactin. *Mol. Biol. Cell* **11,** 3495–3508.

Shea, T. B. (2000). Microtubule motors, phosphorylation and axonal transport of neurofilaments. *J. Neurocytol.* **29,** 873–887.

Shi, J., and Sugrue, S. P. (2000). Dissection of protein linkage between keratins and pinin, a protein with dual location at desmosome-intermediate filament complex and in the nucleus. *J. Biol. Chem.* **275,** 14910–14915.

Simon, M., and Green, H. (1984). Participation of membrane-associated proteins in the formation of the cross-linked envelope of the keratinocyte. *Cell* **36,** 827–834.

Simon, M., Haftek, M., Sebbag, M., Montezin, M., Girbal-Neuhauser, E., Schmitt, D., and Serre, G. (1996). Evidence that filaggrin is a component of cornified cell envelopes in human plantar epidermis. *Biochem. J.* **317,** 173–177.

Sjuve, R., Arner, A., Li, Z., Mies, B., Paulin, D., Schmittner, M., and Small, J. V. (1998). Mechanical alterations in smooth muscle from mice lacking desmin. *J. Muscle Res. Cell Motil.* **19,** 415–429.

Skalli, O., Jones, J. C. R., Gagescu, R., and Goldman, R. D. (1994). IFAP 300 is common to desmosomes and hemidesmosomes and is a possible linker of intermediate filaments to these junctions. *J. Biol. Chem.* **125,** 159–170.

Skerrow, C. J., and Matoltsy, A. G. (1974). Chemical characterization of isolated epidermal desmosomes. *J. Biol. Chem.* **63,** 524–530.

Small, J. V. (1995). Structure-function relationships in smooth muscle: The missing links. *Bioessays* **17,** 785–792.

Small, J. V., Furst, D. O., and Thornell, L.-E. (1992). The cytoskeletal lattice of muscle cells. *Eur. J. Biochem.* **208,** 559–572.

Small, J. V., and Gimona, M. (1998). The cytoskeleton of the vertebrate smooth muscle cell. *Acta Physiol. Scand.* **164,** 341–348.

Smith, E. A., and Fuchs, E. (1998). Defining the interactions between intermediate filaments and desmosomes. *J. Biol. Chem.* **141,** 1229–1241.

Spazierer, D., Fuchs, P., Proll, V., Janda, L., Oehler, S., Fischer, I., Hauptmann, R., and Wiche, G. (2003). Epiplakin gene analysis in mouse reveals a single exon encoding a 725-kDa protein with expression restricted to epithelial tissues. *J. Biol. Chem.* **278,** 31657–31666.

Stanley, J. R., Hawley-Nelson, P., Yuspa, S. H., Shevach, E. M., and Katz, S. I. (1981). Characterization of bullous pemphigoid antigen: A unique basement membrane protein of stratified squamous epithelia. *Cell* **24,** 897–903.

Stappenbeck, T. S., Bornslaeger, E. A., Corcoran, C. M., Luu, H. H., Virata, M. L. A., and Green, K. J. (1993). Functional analysis of desmoplakin domains: Specification of the interaction with keratin versus vimentin intermediate filament networks. *J. Biol. Chem.* **123,** 691–705.

Stappenbeck, T. S., and Green, K. J. (1992). The desmoplakin carboxyl terminus coaligns with and specifically disrupts intermediate filament networks when expressed in cultured cells. *J. Biol. Chem.* **116,** 1197–1209.

Stappenbeck, T. S., Lamb, J. A., Corcoran, C. M., and Green, K. J. (1994). Phosphorylation of the desmoplakin COOH terminus negatively regulates its interaction with keratin intermediate filament networks. *J. Biol. Chem.* **269,** 29351–29354.

Steinert, P. M., Cantieri, J. S., Teller, D. C., Lonsdale-Eccles, J. D., and Dale, B. A. (1981). Characterization of a class of cationic proteins that specifically interact with intermediate filaments. *Proc. Natl. Acad. Sci. USA* **78,** 4097–4101.

Steinert, P. M., Chou, Y. H., Prahlad, V., Parry, D. A. D., Marekov, L. N., Wu, K. C., Jang, S. I., and Goldman, R. D. (1999). A high molecular weight intermediate filament-associated protein in BHK-21 cells is nestin, a type VI intermediate filament protein. Limited co-assembly *in vitro* to form heteropolymers with type III vimentin and type IV alpha-internexin. *J. Biol. Chem.* **274,** 9881–9890.

Steinert, P. M., and Marekov, L. N. (1995). The proteins elafin, filaggrin, keratin intermediate filaments, loricrin, and small proline-rich proteins 1 and 2 are isodipeptide cross-linked components of the human epidermal cornified cell envelope. *J. Biol. Chem.* **270,** 17702–17711.

Steinert, P. M., and Marekov, L. N. (1999). Initiation of assembly of the cell envelope barrier structure of stratified squamous epithelia. *Mol. Biol. Cell* **10,** 4247–4261.

Steinert, P. M., Parry, D. A. D., and Marekov, L. N. (2003). Trichohyalin mechanically strengthens the hair follicle: Multiple cross-bridging roles in the inner root sheath. *J. Biol. Chem.* **278,** 41409–41419.

Styers, M. L., Salazar, G., Love, R., Peden, A. A., Kowalczyk, A. P., and Faundez, V. (2004). The endo-lysosomal sorting machinery interacts wtih the intermediate filament cytoskeleton. *Mol. Biol. Cell.* **15,** 5369–5382.

Sun, D., Leung, C. L., and Liem, R. K. (2001). Characterization of the microtubule binding domain of microtubule actin crosslinking factor (MACF): Identification of a novel group of microtubule-associated proteins. *J. Cell Sci.* **114,** 161–172.

Suzuki, T., Nakamoto, T., Ogawa, S., Seo, S., Matsumura, T., Tachibana, K., Morimoto, C., and Hirai, H. (2002). MICAL, a novel CasL interacting molecule, associates with vimentin. *J. Biol. Chem.* **277,** 14933–14941.

Svitkina, T. M., Verkhovsky, A. B., and Borisy, G. G. (1996). Plectin sidearms mediate interaction of intermediate filaments with microtubules and other components of the cytoskeleton. *J. Biol. Chem.* **135,** 991–1007.

Tarcsa, E., Marekov, L. N., Andreoli, J., Idler, W. W., Candi, E., Chung, S. I., and Steinert, P. M. (1997). The fate of trichohyalin. Sequential post-translational modifications by peptidyl-arginine deiminase and transglutaminases. *J. Biol. Chem.* **272,** 27893–27901.

Tarcsa, E., Marekov, L. N., Mei, G., Melino, G., Lee, S. C., and Steinert, P. M. (1996). Protein unfolding by peptidylarginine deiminase. Substrate specificity and structural relationships of the natural substrates trichohyalin and filaggrin. *J. Biol. Chem.* **271,** 30709–30716.

Thulin, C. D., and Walsh, K. A. (1995). Identification of the amino terminus of human filaggrin using differential LC/MS techniques: Implications for profilaggrin processing. *Biochemistry* **34,** 8687–8692.

Titeux, M., Brocheriou, V., Xue, Z., Gao, J., Pellissier, J. F., Guicheney, P., Paulin, D., and Li, Z. (2001). Human synemin gene generates splice variants encoding two distinct intermediate filament proteins. *Eur. J. Biochem.* **268,** 6435–6449.

Toivola, D. M., Nieminen, M. I., Hesse, M., He, T., Baribault, H., Magin, T. M., Omary, M. B., and Eriksson, J. E. (2001). Disturbances in hepatic cell-cycle regulation in mice with assembly-deficient keratins 8/18. *Hepatology* **34,** 1174–1183.

Traub, P., Perides, G., Kuhn, S., and Scherbarth, A. (1987). Efficient interaction of nonpolar lipids with intermediate filaments of the vimentin type. *Eur. J. Cell Biol.* **43,** 55–64.

Traub, P., Perides, G., Scherbarth, A., and Traub, U. (1985). Tenacious binding of lipids to vimentin during its isolation and purification from Ehrlich ascites tumor cells. *FEBS Lett.* **193,** 217–221.

Troyanovsky, S. M., Eshkind, L. G., Troyanovsky, R. B., Leube, R. E., and Franke, W. W. (1993). Contributions of cytoplasmic domains of desmosomal cadherins to desmosome assembly and intermediate filament anchorage. *Cell* **72,** 561–574.

Tsukita, S., and Tsukita, S. (1985). Desmocalmin: A calmodulin-binding high molecular weight protein isolated from desmosomes. *J. Biol. Chem.* **101,** 2070–2080.

Tsuruta, D., and Jones, J. C. (2003). The vimentin cytoskeleton regulates focal contact size and adhesion of endothelial cells subjected to shear stress. *J. Cell Sci.* **116,** 4977–4984.

Tzivion, G., Luo, Z. J., and Avruch, J. (2000). Calyculin A-induced vimentin phosphorylation sequesters 14-3-3 and displaces other 14-3-3 partners *in vivo*. *J. Biol. Chem.* **275,** 29772–29778.

Uitto, J., Pulkkinen, L., Smith, F. J. H., and McLean, W. H. I. (1996). Plectin and human genetic disorders of the skin and muscle. *Exp. Dermatol.* **5,** 237–246.

Vasioukhin, V., Bowers, E., Bauer, C., Degenstein, L., and Fuchs, E. (2001). Desmoplakin is essential in epidermal sheet formation. *Nat. Cell Biol.* **3,** 1076–1085.

Vicart, P., Caron, A., Guicheney, P., Li, Z., Prevost, M. C., Faure, A., Chateau, D., Chapon, F., Tome, F., Dupret, J. M., Paulin, D., and Fardeau, M. (1998). A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy. *Nat. Genet.* **20,** 92–95.

Vincent, P. A., Xiao, K., Buckley, K. M., and Kowalczyk, A. P. (2004). VE-cadherin: Adhesion at arm's length. *Am. J. Physiol. Cell Physiol.* **286,** C987–997.

Virata, M. L. A., Wagner, R. M., Parry, D. A. D., and Green, K. J. (1992). Molecular structure of the human desmoplakin I and II amino terminus. *Proc. Natl. Acad. Sci.* **89,** 544–548.

Wang, L., Ho, C. L., Sun, D., Liem, R. K., and Brown, A. (2000). Rapid movement of axonal neurofilaments interrupted by prolonged pauses. *Nat. Cell Biol.* **2,** 137–141.

Wang, P., and Gusev, N. B. (1996). Interaction of smooth muscle calponin and desmin. *FEBS Lett.* **392,** 255–258.

Weber, K. L., and Bement, W. M. (2002). F-actin serves as a template for cytokeratin organization in cell free extracts. *J. Cell Sci.* **115,** 1373–1382.

Westgate, G. E., Weaver, A. C., and Couchman, J. R. (1985). Bullous pemphigoid antigen localization suggests an intracellular association with hemidesmosomes. *J. Invest. Dermatol.* **84,** 218–224.

Wiche, G. (1989). Plectin: General overview and appraisal of its potential role as a subunit protein of the cytomatrix. *Crit. Rev. Biochem. Mol. Biol.* **24,** 41–67.

Wiche, G. (1998). Role of plectin in cytoskeleton organization and dynamics. *J. Cell Sci.* **111,** 2477–2486.

Wiche, G., Becker, B., Luber, K., Weitzer, G., Castanon, M. J., Hauptmann, R., Stratowa, C., and Stewart, M. (1991). Cloning and sequencing of rat plectin indicates a 466-kD polypeptide chain with a three-domain structure based on a central alpha-helical coiled coil. *J. Biol. Chem.* **114,** 83–99.

Wiche, G., Herrmann, H., Leichtfried, F., and Pytela, R. (1982). Plectin: A high-molecular-weight cytoskeletal polypeptide component that copurifies with intermediate filaments of the vimentin type. *Cold Spring Harb. Symp. Quant. Biol.* **46,** 475–482.

Wiche, G., Krepler, R., Artlieb, U., Pytela, R., and Aberer, W. (1984). Identification of plectin in different human cell types and immunolocalization at epithelial basal cell surface membranes. *Exp. Cell Res.* **155,** 43–49.

Wiche, G., Krepler, R., Artlieb, U., Pytela, R., and Denk, H. (1983). Occurrence and immunolocalization of plectin in tissues. *J. Biol. Chem.* **97,** 887–901.

Wilson, P. D., Geng, L., Li, X., and Burrow, C. R. (1999). The PKD1 gene product, ''polycystin-1,'' is a tyrosine-phosphorylated protein that colocalizes with alpha2beta1-integrin in focal clusters in adherent renal epithelia. *Lab. Invest.* **79,** 1311–1323.

Xia, C. H., Roberts, E. A., Her, L. S., Liu, X., Williams, D. S., Cleveland, D. W., and Goldstein, L. S. (2003). Abnormal neurofilament transport caused by targeted disruption of neuronal kinesin heavy chain KIF5A. *J. Cell Biol.* **161,** 55–66.

Xu, G. M., Sikaneta, T., Sullivan, B. M., Zhang, Q., Andreucci, M., Stehle, T., Drummond, I., and Arnaout, M. A. (2001). Polycystin-1 interacts with intermediate filaments. *J. Biol. Chem.* **276,** 46544–46552.

Xue, Z. G., Cheraud, Y., Brocheriou, V., Izmiryan, A., Titeux, M., Paulin, D., and Li, Z. (2004). The mouse synemin gene encodes three intermediate filament proteins generated by alternative exon usage and different open reading frames. *Exp. Cell Res.* **298,** 431–444.

Yabe, J. T., Jung, C., Chan, W. K., and Shea, T. B. (2000). Phospho-dependent association of neurofilament proteins with kinesin *in situ*. *Cell Motil. Cytoskeleton* **45,** 249–262.

Yabe, J. T., Pimenta, A., and Shea, T. B. (1999). Kinesin-mediated transport of neurofilament protein oligomers in growing axons. *J. Cell Sci.* **112**(Pt 21), 3799–3814.

Yamazaki, M., Ishidoh, K., Suga, Y., Saido, T. C., Kawashima, S., Suzuki, K., Kominami, E., and Ogawa, H. (1997). Cytoplasmic processing of human profilaggrin by active mu-calpain. *Biochem. Biophys. Res. Commun.* **235,** 652–656.

Yang, H.-Y., Lieska, N., Goldman, A., and Goldman, B. (1985). A 300,000 MW intermediate filament-associated protein in baby hamster kidney (BHK-21) cells. *J. Biol. Chem.* **100,** 620–631.

Yang, Y., Bauer, C., Strasser, G., Wollman, R., Julien, J. P., and Fuchs, E. (1999). Integrators of the cytoskeleton that stabilize microtubules. *Cell* **98,** 229–238.

Yang, Y., Dowling, J., Yu, Q.-C., Kouklis, P., Cleveland, D. W., and Fuchs, E. (1996). An essential cytoskeletal linker protein connecting actin microfilaments to intermediate filaments. *Cell* **86,** 655–665.

Yoon, K. H., Yoon, M., Moir, R. D., Khuon, S., Flitney, F. W., and Goldman, R. D. (2001). Insights into the dynamic properties of keratin intermediate filaments in living epithelial cells. *J. Cell Biol.* **153,** 503–516.

Yoon, M., Moir, R. D., Prahlad, V., and Goldman, R. D. (1998). Motile properties of vimentin intermediate filament networks in living cells. *J. Cell Biol.* **143,** 147–157.

Zamir, E., and Geiger, B. (2001). Molecular complexity and dynamics of cell-matrix adhesions. *J. Cell Sci.* **114,** 3583–3590.

Zastrow, M. S., Vlcek, S., and Wilson, K. L. (2004). Proteins that bind A-type lamins: Integrating isolated clues. *J. Cell Sci.* **117,** 979–987.

Zatloukal, K., Stumptner, C., Lehner, M., Denk, H., Baribault, H., Eshkind, L. G., and Franke, W. W. (2000). Cytokeratin 8 protects from hepatotoxicity, and its ratio to cytokeratin 18 determines the ability of hepatocytes to form Mallory bodies. *Am. J. Pathol.* **156,** 1263–1274.

Zhurinsky, J., Shtutman, M., and Ben-Ze'ev, A. (2000). Plakoglobin and β-catenin: Protein interactions, regulation and biological roles. *J. Cell Sci.* **113,** 3127–3139.

SPECTRIN, α-ACTININ, AND DYSTROPHIN

By M. J. F. BRODERICK* AND S. J. WINDER†

*Department of Biomedical Science, University of Sheffield, Sheffield S10 2TN, United Kingdom;
†Department of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences,
University of Glasgow, Glasgow G12 8QQ, United Kingdom

I. Introduction 204
II. Evolution 205
III. Structure 208
IV. Function 210
A. Actin-Binding Domains 213
B. Spectrin Repeat Region 217
C. Other Binding Partners 220
V. Regulations of Interactions 221
A. CH Domains 221
B. EF Hands 222
C. Lipid Binding 224
D. Polyproline Binding Domains 225
E. ZZ domain 227
VI. Disease 227
A. Duchenne Muscular Dystrophy 227
B. Hereditary Spherocytosis 229
C. Familial Focal Segmental Glomerulosclerosis 231
VII. Conclusion 232
References 233

Abstract

Spectrin family proteins represent an important group of actin-bundling and membrane-anchoring proteins found in diverse structures from yeast to man. Arising from a common ancestral α-actinin gene through duplications and rearrangements, the family has increased to include the spectrins and dystrophin/utrophin. The spectrin family is characterized by the presence of spectrin repeats, actin binding domains, and EF hands. With increasing divergence, new domains and functions have been added such that spectrin and dystrophin also contain specialized protein-protein interaction motifs and regions for interaction with membranes and phospholipids. The acquisition of new domains also increased the functional complexity of the family such that the proteins perform a range of tasks way beyond the simple bundling of actin filaments by α-actinin in *S. pombe*. We discuss the evolutionary, structural, functional, and regulatory roles of the spectrin family of proteins and describe some of the disease traits associated with loss of spectrin family protein function.

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70007-8

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

I. Introduction

The eukaryotic cytoskeletal network is formed from a number of filamentous systems composed of polymers of actin, tubulin, or intermediate filament proteins. The actin stress fibers, microtubules, and intermediate filaments generated are integrated in a highly organized manner that can be both dynamic as well as stable. The filamentous state and organization of these proteins provides the cell with an internal scaffold essential to many cellular processes, including mechanical strength, cellular morphology, adhesion, motility, intracellular trafficking, cell division, and networks for inter- and intracellular communication. The cytoskeletal network allows rapid remodeling in response to altered mechanical needs facilitated by the dynamic exchange of protein subunits within the system, and by the manner in which the network is linked through crosslinking proteins.

Significant contributors to cell structure are those proteins that crosslink actin filaments or connect actin filaments to the cell membrane. Examples of such proteins can be found within the spectrin superfamily of cytoskeletal proteins. This discrete group is principally composed of the actin crosslinking protein α-actinin, and the membrane-associated actin-binding proteins spectrin and dystrophin.

The spectrin family of proteins is highly modular and share common structural elements including a calponin homology (CH) domain containing an actin-binding domain, spectrin repeats, EF hands, and various other signaling domains and motifs (Fig. 1). The spectrin family of proteins arose from work that originally focused on understanding the role that spectrin played biochemically in the organization and assembly of the cytoskeleton (reviewed in Gratzer, 1985). Spectrin possesses the ability to self assemble, but the molecular basis of this process could not be explained until more was known about the sequence and, ultimately, the structure of the protein. The work of Speicher and Marchesi (1984) provided the protein sequence of almost half of the α-spectrin chain. This work identified spectrin as being a highly modular protein composed of many repeating 106-amino-acid units. The helical nature of these units was predicted to form triple-helical coiled-coil bundles that were dubbed *spectrin repeats*. Continued investigation and DNA sequencing led to the determination of several α-spectrin sequences from erythrocyte, *Drosophila*, and brain (Dubreuil *et al.*, 1989; Sahr *et al.*, 1990; Wasenius *et al.*, 1989). It was around this time that the DNA sequences of the related proteins α-actinin and dystrophin were completed (Baron *et al.*, 1987; Koenig *et al.*, 1988). These proteins were found to contain repeating units similar to those found in spectrin (Davison and Critchley, 1988) and, hence, the spectrin family of proteins was born.

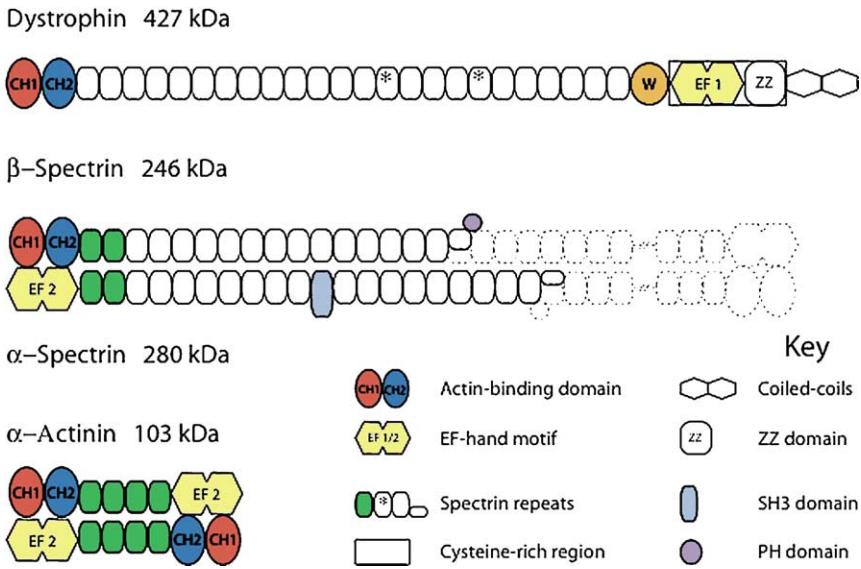


FIG. 1. Structure of spectrin superfamily proteins. Modular domains within each protein are clearly defined. Shaded spectrin repeats represent coiled coils involved in dimerization events; incomplete repeats represent proportionally the number of coiled-coil helices contributed by α- and β-spectrin when generating a complete spectrin repeat during formation of the spectrin tetramer. The dashed lines indicate how two spectrin heterodimers interact to form a functional spectrin tetramer. Asterisks in the dystrophin spectrin repeats represent the position of the two greater repeats in dystrophin with respect to utrophin, which in all other respects has a similar overall structure. Numbers in the EF hand regions represent the number of EF hand motifs.

The sequencing of α- and β-spectrin, α-actinin, and dystrophin has revealed similarities not only within the spectrin repeat, but also the other domains and motifs present within these proteins. Subsequent analyses have revealed an evolutionary pathway for the divergence of spectrin and dystrophin from a common α-actinin ancestor via a series of rearrangements, duplications, and evolution of repeats and other domains, as well as the acquisition of unique domains such as PH, WW, and SH3 (Fig. 2).

II. EVOLUTION

The availability of complete sequences for α-actinin, spectrin, and dystrophin has allowed the ancestry and evolution of the proteins to be traced. Multiple sequence alignments and phylogenetic trees have been

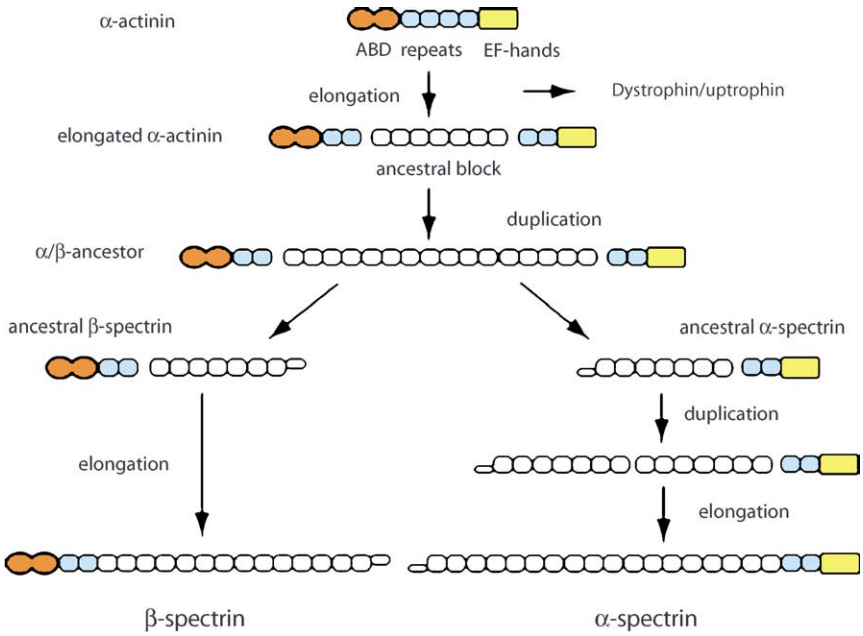


FIG. 2. Evolution of the spectrin superfamily. Rounded rectangles represent spectrin repeats. Shaded rectangles denote α-actinin-like repeats involved in dimerization, whereas unshaded rectangles represent repeats that were involved in duplication and/or elongation events. The incomplete spectrin repeats involved in tetramer formation are proportionally represented depending on the number of repeat helices each protein contributes to the formation of a complete spectrin repeat. (Adapted from Dubreuil, 1991; Pascual *et al.*, 1997.) A dystrophin/utrophin ancestor probably diverged from α-actinin at a relatively early stage and then underwent its own series of duplications and acquisitions of new motifs.

combined with precise alignment of equivalent domains within each protein to provide details of the relationships that exist between each of the protein domains. Analysis of the amino acid sequences from α-actinin, spectrin, and dystrophin suggested that all three of the protein families have arisen from a common ancestral protein that was similar to α-actinin (Byers *et al.*, 1992; Dubreuil, 1991) via a series of gene duplications and gene rearrangements (Baines, 2003; Pascual *et al.*, 1997; Thomas *et al.*, 1997; Viel, 1999). One of the diagnostic features of the spectrin superfamily of proteins is the presence of the 106–120 repetitive unit referred to as the *spectrin repeat* (Speicher and Marchesi, 1984). Sequence comparisons suggest that all three of these proteins share significant homology within their

N-terminal actin-binding domains and in the spectrin repeats that form the rod domains (Davison and Critchley, 1988). The spectrin repeats are found in distinct multiples in each protein, resulting in a characteristic actin crosslinking distance. α-Actinin contains four repeats, β-spectrin contains 17, α-spectrin contains 20, and dystrophin contains 24. The sequences of some spectrin repeats of α- and β-spectrin are similar in many ways to the four repeats present in α-actinin (Dubreuil, 1991). Within the cell, α-actinin and spectrin dimerize, although the spectrins interact further to generate a functional tetramer (Fig. 1). Most notable is that the ends of the native spectrin tetramer involved in the dimerization event show remarkable similarity to the rod domain repeats of α-actinin that also mediate dimer formation.

Indeed, homologous regions of all of the α-actinin protein domains can be found within the sequences of α- and β-spectrin. For example, the amino and carboxy terminal regions of α-actinin resemble the N-terminus of β-spectrin and the C-terminus of α-spectrin, respectively (Byers *et al.*, 1989; Dubreuil *et al.*, 1989). Phylogenetic analysis shows a common ancestor for the first repeat of α-actinin and the first repeat of β-spectrin. Similarly, each of the remaining repeats in α-actinin (2–4) correspond to repeats 1 and 2 of β-spectrin and repeats 19 and 20 of α-spectrin respectively (Fig. 2). This may have relevance for the function of these repeats in the dimerization of these proteins (Pascual *et al.*, 1997). It is the similarity between these regions of α-actinin, the spectrins, and the simpler domain organization of α-actinin that have led to the hypothesis that these two protein families have evolved from an α-actinin-like precursor.

Spectrin is a much more elongated protein compared to α-actinin due to the additional number of repeats. The additional repeats are more closely related to one another than repeats common to both α-actinin and spectrin. The spectrin repeat sequences are the most divergent in dystrophin and its homologue utrophin (Winder *et al.*, 1995a), most likely reflecting an earlier divergent event when compared to spectrin (Pascual *et al.*, 1997).

The additional molecular length of spectrin compared to the ancestral α-actinin is believed to have arisen through two major duplication events containing blocks of seven repeats (Pascual *et al*, 1997). The beginning of the process can be described as the elongation of α-actinin by insertion of a seven-repeat block between the second and third α-actinin repeats. That block of repeats was then duplicated and an ancestor of the tetramerization repeat was inserted between the blocks (Pascual *et al.*, 1997) (Fig.2). Overall, the two-stage evolution of the superfamily is believed to have involved an initial dynamic phase involving intragenic duplications and concerted evolution, followed by a stable phase where repeat numbers

became constant and their sequences evolved independently (Thomas *et al.*, 1997). The first phase of evolution resulted in the α-actinin, spectrin, and dystrophin lineages via a series of duplication events. This process gave rise to the necessary repeat lengths and numbers within the spectrin and dystrophin lineages. The α-actinin lineage continued to evolve. Phylogenetic analysis has indicated that the different isoforms found today in modern vertebrates arose after the vertebrate urochordate split, with the muscle and nonmuscle isoforms evolving separately (Virel and Backman, 2004). Thomas and colleagues (1997) hypothesized that, during a transition period, the new genes began to acquire distinct crosslinking distances and that subsequent selection against longer or shorter proteins would result in a stabilization of protein length. The current length of spectrin repeats is evidently very stable, as there has been little change since the split of the arthropod vertebrate lineages.

III. Structure

The spectrin superfamily is an important group of cytoskeletal proteins involved in many functions requiring crosslinking, bundling, or binding to filamentous actin. Each of the proteins within this family differs greatly in its specific biological function. However, they all share a surprising level of structural homology. The members of this protein family are composed of a number of conserved domains: spectrin repeats, CH domain containing actin-binding domains, EF hands, calcium-binding motifs, and various signaling domains (Fig. 1). The actin-binding domain (ABD) is the most N-terminal domain and can be found in α-actinin, β-spectrin, dystrophin, and utrophin, but not α-spectrin. The presence of the ABD allows these proteins to interact with F-actin in a variety of different cellular situations.

The ABDs of these proteins comprise a tandem pair of CH domains, although the manner in which this domain interacts with F-actin is still unclear even after extensive investigation and modeling. The spectrin repeats form the rod domains of these proteins. There are four repeats in α-actinin and between 17 and 24 in spectrin and dystrophin. In general, the spectrin repeats are responsible for the overall length of the protein and they are usually recognized as modules for the generation of elongated molecules and the separation of the specific N- and C-terminal domains (Winder, 1997). Overall, the core structures of the spectrin repeats from each family member are very similar, although the size of the repeating regions differs slightly. In α-actinin, repeats are 122 residues in length, and 106 residues in length for spectrin; in dystrophin and utrophin, the repeating units are 109 residues in length. Within the spectrin family proteins, each member contains a differing number of repeating

elements. The four repeats in α-actinin are separated from the ABD by a linker that allows a significant degree of flexibility between the rigid rod and the ABD. In the functional α-actinin dimer, it is the spectrin repeats that are responsible for the dimerization. Crystal structures of the dimeric α-actinin rod domain have found that the rod bends slightly along its length, but also the whole domain twists through approximately 90 degrees. In conjunction with the rod domain and ABD, the flexible linker that separates them contributes to α-actinin's ability to crosslink actin filaments that are oriented in either a parallel or antiparallel manner (Ylanne *et al.*, 2001a).

The rod domains of α- and β-spectrin are composed of 20 and 17 spectrin repeats, respectively. At the N-terminus of α-spectrin, the first repeating segment begins with the ''third helix'' of what will become a complete triple-helical coiled-coil structure, analogous to a spectrin repeat when it interacts with the two helices that are found at the C-terminus of β-spectrin (Fig. 1). This site allows the spectrin dimer to interact with another dimer to form a spectrin heterodimer or tetramer. This is the functional unit of spectrin within the erythrocyte. It has been shown that specific repeats in both α- and β-spectrin are not just present as structural modules that contribute to the length of the rod domain. Repeat 10 of α-spectrin has been found to be slightly shorter than a typical spectrin repeat and shows substantial homology to the SH3 domain of the Src protein family (Wasenius *et al.*, 1989), whereas repeat 15 of β-spectrin has been found to be responsible for interaction with ankyrin (Kennedy *et al.*, 1991). The two most C-terminal spectrin repeats of α-spectrin (21 and 22) and the first two of β-spectrin (1 and 2) once again show differences in sequence and structure from a typical spectrin repeat. This particular section of α-spectrin shows homology with the C-terminus of α-actinin, whereas repeats 1 and 2 of β-spectrin share homology with the N-terminus. It has been found that these repeats are responsible for the dimerization of α- and β-spectrin in a manner analogous to the four spectrin repeats of the α-actinin rod domain (reviewed in Winkelmann and Forget, 1993). The rod domains of dystrophin and its close relative utrophin have not been found to mediate dimerization. It seems that the spectrin repeats of these two proteins function primarily to separate the N- and C-termini. However, it should be noted that the rod domains of dystrophin and utrophin are able to associate with filamentous actin, although the manner of interaction differs for each protein (Amann *et al*, 1999; Rybakova *et al.*, 2002).

The C-termini of spectrin family proteins also exhibit a variety of structural similarities including motifs involved in protein-protein interactions (Fig. 1). EF-hands motifs are found in all but β-spectrin. The

function of this domain is most notable in α-actinin, where the binding of calcium is able to affect the interaction of the ABD with F-actin. This is only the case though for the nonmuscle isoforms of α-actinin, as the muscle isoforms are calcium insensitive (Blanchard *et al.*, 1989). The C-terminus of α-spectrin has also been found to contain EF hands similar to those of α-actinin. It is believed that these structures may play a role in modulating the functional conformation of the β-spectrin ABD when N- and C-termini are juxtaposed in the spectrin heterodimers. The EF hands of dystrophin have been predicted to be unable to bind calcium, although it is thought that they have an important structural role (Huang *et al.*, 2000). Finally, a number of smaller motifs have been found at or within the C-termini of these proteins. The nonerythroid form of β-spectrin has been shown to possess a pleckstrin homology (PH) domain at its C-terminus. This fold is conserved in proteins that can interact with phospholipids and may allow direct interaction with the cell membrane. The remainder of the motifs are found within the C-terminus of dystrophin and utrophin. These motifs consist of WW and ZZ domains that, together with the EF-hand region, form the cysteine-rich region. The WW domain is an example of a protein-protein interaction module that binds proline-rich sequences (Ilsley *et al.*, 2002), whereas the ZZ domain is a zinc finger motif also involved in mediating protein-protein interactions (Ponting *et al.*, 1996). At the extreme C-terminus of dystrophin and utrophin are two helices predicted to form dimeric coiled-coils (Blake *et al.*, 1995) that mediate interaction with the dystrophin family proteins dystrobrevin and DRP2.

IV. Function

Spectrin is a common component of the submembranous cytoskeleton. It was first identified as a major constituent of the erythrocyte membrane cytoskeleton, but has since been found in many other vertebrate tissues as well as in the nonvertebrates *Drosophila, Acanthamoeba, Dictyostelium*, and echinoderms (Bennett and Condeelis, 1988; Byers *et al.*, 1992; Dubreuil *et al.*, 1989; Pollard, 1984; Wessel and Chen., 1993). The ease with which spectrin could be isolated from erythrocyte ghosts made it an ideal candidate for the study of the biochemical processes involved in the assembly and organization of the cytoskeleton (Gratzer, 1985).

The human erythrocyte possesses a characteristic biconcave shape and remarkable viscoelastic properties. Electron microscopy studies performed on red blood cells (RBC), ghosts, and skeletons revealed a two-dimensional lattice of cytoskeletal proteins. This meshwork of proteins was thought to determine the elastic properties of the RBC. This

notion was supported further when it was found that detergent-extracted skeletons exhibited shape memory; since spectrin was the major constituent of these skeletons, it was suggested that it was responsible for the elastic properties. The protein lattice that laminates the inner surface of the erythrocyte membrane is formed from interactions between actin, spectrin, and integral membrane proteins (Bennett and Gilligan, 1993). The lattice is predominantly formed from α- and β-spectrin dimers that again dimerize to generate tetramers roughly 200 nm in length. Five or six of these tetramers bind through their tail ends to a junctional complex, consisting of filamentous actin and band 4.1 (reviewed in Winkelmann and Forget, 1993). The molecular function of the complete spectrin heterodimer relies on the inter- and intramolecular interactions that occur at two key points within the spectrin molecule. These associations take place at the head end, where interchain binding between α- and β-spectrin gives rise to a heterodimer or the tetramer. The tail end of the molecule contains sites responsible for the interchain binding between spectrin chains integrating spectrin tetramers into a network via interactions with actin, protein 4.1, and other binding partners. Between the head and tail regions of the molecule, much of the overall length of spectrin is attributed to the number of spectrin repeats, 20 repeats for α-spectrin and 17 repeats for β-spectrin. The spectrin tetramers associate with short actin oligomers to form a regular repeating polygonal lattice. This network is coupled to the membrane via a limited number of direct and indirect contacts between spectrin and integral membrane proteins. These attachments consist of interactions between ankyrin and band 3 protein, and between protein 4.1 and glycophorin C.

α-Actinin is the smallest member of the spectrin family of proteins (Pascual *et al.*, 1997). It was first described as an actin crosslinker in skeletal muscle, but has subsequently been found to be ubiquitously expressed (Otto, 1994). Additional family members have been found in smooth muscle and nonmuscle cells (Blanchard *et al.*, 1989) and are localized at the leading edge, cell adhesion sites, focal contacts, and along actin-stress fibers in migrating cells (Knight *et al.*, 2000). The functional unit of α-actinin is an antiparallel homodimer of polypeptide chain mass of 94–103 kDa (Blanchard *et al.*, 1989), in which the amino-terminal CH domains together with the carboxy-terminal calmodulin (CaM) homology domain form the actin-binding heads of the molecule (Critchley, 2000). The connection between these two heads is composed of four spectrin repeats, which define the distance between the actin filaments that are crosslinked. α-Actinin has three main biological functions. It is the major thin filament crosslinking protein in the muscle Z-discs, where it holds the adjacent sarcomeres together (Masaki *et al.*, 1967). α-Actinin is

also found close to the plasma membrane where it crosslinks cortical actin to integrins (Otey *et al.*, 1990) and serves as a linker between transmembrane receptors and the cytoskeleton. In nonmuscle cells, α-actinin is a major component of stress fibers, an analogous contractile structure to the more organized units found in striated muscle (reviewed by Otey and Carpen, 2004). α-Actinin is also a constituent of dense bodies (Lazarides and Burridge, 1975), which are believed to be structurally and functionally analogous to the sarcomere Z-disk.

Dystrophin is the product of the largest known gene within the human genome, spanning approximately 2.5 Mb of genomic sequence and composed of 79 exons (Coffey *et al.*, 1992; Roberts *et al.*, 1993). The protein product encoded by the transcript of this gene is known as dystrophin, and the absence of this protein results in Duchenne muscular dystrophy (Koenig *et al.*, 1987). Dystrophin is predominantly expressed in skeletal and cardiac muscle but small amounts are found in the brain. These full-length isoforms are under the control of three independently regulated promoters referred to as brain, muscle, and Purkinje, the names of which reflect the site at which dystrophin expression is driven. Additionally, four internal promoters give rise to truncated C-terminal isoforms, and alternative splicing further increases the number of isoforms and variants. The spectrin repeats form the bulk of the protein (Fig. 1) and are thought to allow flexibility and give the molecule a rodlike structure. Dystrophin can be found associated with the plasma membrane of cardiac and skeletal muscle, where it interacts with the integral membrane protein dystroglycan that binds to laminin on the extracellular face. The dystrophin-dystroglycan complex further interacts with the integral membrane sarcoglycan proteins and peripheral membrane proteins syntrophin and dystrobrevin, which together comprise the dystrophin glycoprotein complex (reviewed in Winder *et al.*, 1995a). This complex of proteins can then interact with F-actin via the N-terminus of dystrophin to form a flexible link between the basal lamina of the extracellular matrix and the internal cytoskeletal network (Campbell and Kahl, 1989; Rando, 2001). It is believed that this complex serves to stabilize the sarcolemma and protect muscle fibers from contraction-induced damage. Indeed, the absence or mutation of dystrophin results in the X-linked myopathies Duchenne and Becker muscular dystrophies (DMD and BMD, respectively; reviewed in Blake *et al.*, 2002).

Within the skeletal musculature, dystrophin plays an important role in maintaining the integrity of the sarcolemmal membrane. Dystrophin is not able to perform this task alone and interacts with a number of other proteins that include dystroglycans, sarcoglycans, dystrobrevins, syntrophins, and sarcospan (Straub and Campbell, 1997). Mutation of dystrophin or, indeed,

other components of this complex can result in a variety of disease pathologies. Most notable is that of DMD, where a complete absence of dystrophin is observed, resulting in progressive muscle wasting and eventual death of the affected individual. This complex of proteins is thought to provide a link between the cortical actin cytoskeletal network and laminin in the extracellular matrix (Ervasti and Campbell, 1993a,b). Hence, it is thought that dystrophin and the associated proteins provide a mechanically stabilizing role that protects the sarcolemmal membrane from the shear stresses generated during eccentric contraction of muscle (Petrof *et al.*, 1993). Dystrophin has been found to localize adjacent to the cytoplasmic face of the sarcolemmal membrane in regions known as costameres (Porter *et al.*, 1992; Straub *et al.*, 1992). These assemblies of cytoskeletal proteins are involved in linking the force-generating sarcomeric apparatus to the sarcolemmal membrane (Craig and Pardo, 1983; Pardo *et al.*, 1983). Costameres transmit contractile forces laterally through the sarcolemmal membrane to the basal lamina (Danowski *et al.*, 1992). It has been found that dystrophin is not required for the assembly of several of the proteins that comprise costamere-like structures, but its absence does lead to an altered costameric lattice (Ehmer *et al.*, 1997; Minetti *et al.*, 1992; Porter *et al.*, 1992; Williams and Bloch, 1999). These data suggest that dystrophin plays an important role in the organization or stability of costameres, perhaps via an interaction with actin filaments (Rybakova *et al.*, 2000). Rybakova and colleagues (2000) showed that the dystrophin complex formed a mechanically strong link between the sarcolemma and the costameric cytoskeleton through interaction with γ-actin filaments.

Following the discovery of the dystrophin gene, another cDNA was identified that showed considerable homology to that of dystrophin (Love *et al.*, 1989). Initially this protein was referred to as dystrophin-related protein (DRP), but once cloned and sequenced (Tinsley *et al.*, 1992) it was subsequently renamed utrophin due to a ubiquitous expression pattern compared to that of dystrophin. In muscle cells, utrophin shares a high degree of functional similarity with dystrophin (Claudepierre *et al.*, 1999; Earnest *et al.*, 1995; Loh *et al.*, 2000; Matsumura *et al.*, 1993; Nguyen *et al.*, 1991; Pons *et al.*, 1994; Raats *et al.*, 2000) and has been proposed as a potential therapeutic replacement for dystrophin in the treatment of DMD (Matsumura *et al.*, 1992; Pearce *et al.*, 1993; Tinsley *et al.*, 1992).

A. *Actin-Binding Domains*

The actin-binding domains (ABD) of spectrin, α-actinin, and dystrophin consist of approximately 240 residues that comprise two functionally distinct but structurally equivalent domains (Gimona and Winder, 1998;

Matsudaira, 1991). These domains have been named *calponin homology* or *CH domains* (Castresana and Saraste, 1995) based on the sequence similarity to the smooth muscle regulatory protein calponin (Winder and Walsh, 1990), where this domain is found only as a single copy. Compared to the single CH domain seen in calponin, the double domain found in the spectrin superfamily of proteins has been proposed to arise through a process of gene duplication (Castresana and Saraste, 1995; Matsudaira, 1991). Furthermore, phylogenetic analysis of CH-domain-containing proteins has revealed that there is a greater similarity between the N-terminal CH domains (CH1) and the C-terminal CH domains (CH2) in any of the classes of actin-binding proteins than between the CH domains found within the same protein (Banuelos *et al.*, 1998; Keep *et al.*, 1999a; Korenbaum and Rivero, 2002; Stradal *et al.*, 1998). Some actin-binding proteins have been found to contain a single CH domain, although the binding interaction requires additional elements since the isolated CH domains from these proteins do not exhibit analogous actin binding when compared to the whole protein (Gimona and Mital, 1998). Functionally, it should be noted that characterization of bacterially expressed CH domains corresponding to the actin-binding domains of α-actinin, dystrophin, and utrophin were found not to be equivalent (Way *et al.*, 1992; Winder *et al.*, 1995b). When the two CH domains are separated and then used in actin-binding experiments, it has been found that only CH1 has the ability to interact with F-actin, although the affinity is reduced. CH2 has little or no intrinsic actin-binding activity, but it is obvious that its presence is functionally important. It has been shown that both of the CH domains are required to achieve the greatest interaction with F-actin and hence, single CH domains are not regarded as actin-binding domains per se (Gimona and Winder, 1998).

The first crystal structures of CH domains were published in 1997. These were the CH2 domain of spectrin (Carugo *et al.*, 1997) and the N-terminal ABD of fimbrin, comprising two CH domains (CH1.1 and CH2.1) (Goldsmith *et al.*, 1997). The crystal structure of the second utrophin CH domain (CH2) was later published by Keep *et al.* (1999a). It was not long, however, before the complete actin-binding domain of utrophin was crystallized (Keep *et al.*, 1999b), followed closely by dystrophin (Norwood *et al.*, 2000). The CH domain is a compact globular domain that appears to show a high degree of structural conservation. Overall, the domain comprises four main α-helices (A, C, E, and G) that are approximately 11 to 18 residues in length and exhibit a roughly parallel orientation. Three shorter helices (B, D, and F) are less regular and form lesser secondary structure elements. The structure can be considered to comprise a number of layers. The core of the domain is formed by a parallel

arrangement of helices C and G, which are then sandwiched between helix E and the N-terminal helix A (Broderick and Winder, 2002).

The crystal structures of CH domains from α-actinin, dystrophin, utrophin, fimbrin, spectrin, and plectin (Garcia-Alvarez *et al.*, 2003) have been solved to date. These structures have given insight into certain aspects of CH domain function, but have also raised many new questions regarding the interaction of these domains with actin. The dimeric organization displayed by the crystallized utrophin and dystrophin ABDs contrasts strongly with that of the α-actinin and the related fimbrin actin-binding domain (K. Djinovic-Carugo, personal communication; Goldsmith *et al.*, 1997). When crystallized, fimbrin and α-actinin do so as compact monomers, where the two CH domains fold back on themselves to form a compact globular structure. These same interfaces are involved in the dimerization seen in utrophin, except it is the CH1 and CH2 domains from separate molecules that interact. The preservation of an interface between two domains of the same or related proteins when in monomeric or oligomeric forms is known as three-dimensional domain swapping (Schlunegger *et al.*, 1997). Recent cryo-EM reconstructions of these domains with F-actin have not served to definitively resolve the mode of interaction of these proteins with F-actin. Reconstructions of fimbrin bound to F-actin have been modeled on a compact conformation (Hanein *et al.*, 1998) and the reconstruction of utrophin with F-actin on an extended conformation (Moores *et al.*, 2000b). Electron diffraction and modeling of the α-actinin molecule bound to F-actin showed that the ABD could be associating as an open bilobed structure (Tang *et al.*, 2001; Taylor and Taylor, 1993). The cryo-EM reconstruction of α-actinin with F-actin (McGough *et al.*, 1994), however, revealed a more globular difference density, suggesting that α-actinin might also associate with F-actin in a manner more analogous to the compact mode of interaction seen in the fimbrin crystal structure. The helical linkers between the two CH domains of these proteins may play an important role in determining the flexibility between the two CH domains and, subsequently, the manner they interact with F-actin. Gel-filtration studies of the utrophin ABD have shown it to be monomeric when in solution (Winder *et al.*, 1995a). Hence, the crystal structures of the ABDs from fimbrin and utrophin may represent two conformational extremes within this class of actin-binding proteins (Keep *et al.*, 1999b). Several modes of interaction with F-actin have been demonstrated in cryo-EM studies with the utrophin ABD (Galkin *et al.*, 2002).

Within the N-terminal, ABDs of three actin-binding sites (ABS) from spectrin superfamily proteins have been delineated (ABS1, ABS2, and ABS3). The first and third ABS have been localized to the α1 helix in

the first and second CH domains, respectively. These sites were originally identified using synthetic peptides derived from dystrophin (Levine *et al.*, 1990, 1992). The second ABS, corresponding to helices α5 and α6 in the CH1 domain, was first identified in the Dictyostelium actin-gelation factor ABP-120 (Bresnick *et al.*, 1990, 1991). ABS2 was later identified in α-actinin using *in vitro* actin-binding studies with glutathione S-transferase (GST) fusion proteins (Kuhlman *et al.*, 1992). It was later recognized that the ABS sequences were not part of a fully folded globular protein and hence residues not normally involved in actin-binding may have been allowed to interact with actin. Structural approaches have begun to shed light on the mechanism by which the spectrin superfamily of proteins interacts with the cytoskeleton. Biochemical work studying the actin-binding sites of α-actinin (Lebart *et al.*, 1990, 1993; McGough *et al.*, 1994; Mimura and Asano, 1987) and dystrophin (Levine *et al.*, 1992) have been successful in identifying actin subdomain 1 as an important binding site for this particular class of proteins. Electron microscope reconstructions of F-actin decorated with α-actinin (McGough *et al.*, 1994; Tang *et al.*, 2001) reveal that actin subdomain 1 forms the major site of interaction. Helical reconstruction also revealed that these two proteins interacted with adjacent actin monomers on the long pitch helix, a site apparently shared by most F-actin binding proteins (McGough, 1998).

The crystal structure of the utrophin ABD suggested an alternate model to that of the association of fimbrin with actin (Keep *et al.*, 1999b). As utrophin crystallized as a head-to-tail dimer, each of the monomers adopted an extended conformation. This arrangement placed the predicted ABSs on the surface of the protein, clearly enabling interaction with actin. The dimerization of utrophin seen in the crystal conserved the inter-CH domain interfaces, suggesting that utrophin may adopt a more compact conformation when in solution. To date, there is no evidence to support anything other than a monomeric conformation of utrophin when in solution, as the binding stoichiometry with actin is 1:1 (Keep *et al.*, 1999b; Moores and Kendrick-Jones, 2000a; Winder, 1996). However, the crystallization of utrophin as a dimer suggested that the ABDs of this protein might be flexible and allow actin-binding in an open conformation, even when utrophin exists as a monomer in solution. Moores *et al.* (2000b) developed this idea further by demonstrating a model of utrophin-actin binding that contrasted that of fimbrin bound to actin. A pseudo-atomic model of utrophin bound to F-actin in an open conformation was detailed, showing that all of the ABSs could be directly involved in the actin interaction. This mode of binding was found to create a different conformational change within actin compared to that caused by fimbrin,

suggesting that an induced fit mechanism involving conformational flexibility of actin and utrophin may be crucial to their interaction (Moores *et al.*, 2000b). The validity of this model has recently been called into question (Galkin *et al.*, 2003), however, and alternative models have been proposed (Galkin *et al.*, 2002). The comparison between the previously published model of fimbrin binding to actin and that of utrophin shows that both proteins possess a totally different mode of interaction even though their ABDs share sequence homology. Such a difference is likely to be related to the overall function of each protein, but the utrophin model has identified an alternate means of actin association within a large family of proteins important to cellular organization. It should be noted that Moores *et al.* (2000b) did not exclude the compact orientation in utrophin actin binding. However, their model relies on inherent protein flexibility. Indeed, models of utrophin ABDs have been generated where association with F-actin is made in a closed compact conformation (Sutherland-Smith *et al.*, 2003). Recent crystallographic and calorimetric studies of the plectin ABD demonstrated that while the two CH domains associate to form a closed conformation in the crystal structure, binding to F-actin induces the open conformation (Garcia-Alvarez *et al.*, 2003). Elucidation of spectrin's interaction with actin at the molecular level, however, has been hampered by an inability to express a functional spectrin ABD in isolation.

B. Spectrin Repeat Region

The rod domain of α-actinin is the shortest within this family of proteins and comprises just four spectrin-like repeats. Given the reduced length of this domain, it is feasible to assume a greater degree of rigidity, especially as the functional unit of α-actinin is a dimer. The spectrin repeats of the rod domain are essential to the dimerization of α-actinin. The association of the rod domains from two monomers leads to a much more stable and less flexible domain overall (Djinovic-Carugo *et al.*, 2002). Spectrin, dystrophin, and the related protein utrophin all contain many more spectrin repeats that seem to play a more direct role in the cellular function of these proteins. Studies performed by Pasternak and colleagues (1995) show the sarcolemma of muscles from the mdx mouse to be four times less stiff than in controls, demonstrating directly that dystrophin and its associated proteins reinforce the stability of the sarcomere. Spectrin forms roughly 5% of the total protein of the erythrocyte and is pivotal to the formation and function of the submembranous skeleton in red blood cells. This erythrocyte plasma membrane possesses remarkable mechanical

properties, more like an elastic semisolid (Evans and Hochmuth, 1977). This allows storage of energy during deformation (e.g., squeezing through a narrow capillary), but allows the erythrocyte to return to its normal shape once the deformation ceases (see review by Bennett and Gilligan, 1993).

The ability of erythrocytes to survive repeated deformation is essential to their physiological function and their prolonged life within the vasculature. Direct evidence has determined that spectrin is the major factor in providing the elastic properties exhibited by the erythrocyte. Studies of erythrocytes from patients suffering from hereditary spherocytosis clearly demonstrate this (Waugh and Agre, 1988), as reduced quantities of spectrin result in a greater extent of clinical severity and a reduction in the force required to deform the affected erythrocytes (Agre *et al.*, 1985, 1986). This perhaps results from an overall effect of the structure of the submembranous lattice on the whole, but the properties of this meshwork of proteins can be linked to the properties of the spectrin repeats found within the rod domain. For many cytoskeletal and adhesion proteins, the ability to survive extension and deformability is pivotal to their role in a cellular environment. Atomic force microscopy (AFM) has been employed to examine the extensibility of spectrin repeats (Rief *et al.*, 1999). These studies have determined that the α-helical spectrin repeat can be forced to unfold in a stochastic one-domain-at-a-time fashion (Rief *et al.*, 1999).

The availability of tandem spectrin repeat structures from nonerythroid α-spectrin (Grum *et al.*, 1999) and the four repeat rod domain of α-actinin (Ylanne *et al.*, 2001a) have shown that individual spectrin repeats should not be considered as such, and that the interspectrin repeat links are actually formed from contiguous helices rather than flexible linkers (Law *et al.*, 2003). This has implications for the manner that spectrin repeats respond to mechanical stress, inasmuch as the repeats within the rod domain do not unfold one at a time. Rather, they are subject to a cooperative manner of forced unfolding (Law *et al.*, 2003). Helical linkers between spectrin repeats have been implicated to help explain the extensibility and elasticity observed within the erythrocyte cytoskeleton. The unfolding of spectrin repeats might explain thermal-softening (Waugh and Evans, 1979) and strain softening of the RBC submembranous network (Lee and Discher, 2001; Markle *et al.*, 1983). Additionally, it has also been shown that tandem spectrin repeats are thermodynamically more stable than individual repeats and that tandem repeats unfold in unison, behaving similarly to an individual repeat (MacDonald and Pozharski, 2001).

The rod domains of spectrin family proteins are assumed to function solely as structural spacers that serve to separate the C-terminus from the N-terminal actin-binding domain. While this is likely to be the case for α-actinin (and to a certain extent in spectrin), natural mutations and transgenic experiments would suggest otherwise for dystrophin. It was widely thought that the rod domains of dystrophin (and utrophin) served as flexible spacers or shock absorbers between the actin cytoskeleton and the sarcolemmal membrane (Winder *et al.*, 1995a). However, is the length of this ''shock absorber'' crucial to the function of the protein? An individual with a large deletion in the dystrophin gene encompassing 46% of the entire protein and 73% of the rod region (repeats 4–19) presented with a very mild BMD phenotype (England *et al.*, 1990). This would tend to suggest that rod domain length is not essential with regards to the proposed shock-absorbing role of the protein. The fact that a dystrophic phenotype is observed, regardless of how mild, would suggest that there is functional importance regarding the length of the rod domain. However, dystrophin minigenes have been designed on the basis of this shortened dystrophin and have been used to correct the dystrophic phenotype in mdx mice (Phelps *et al.*, 1995; Wells *et al.*, 1995).

Similarly, a minispectrin has also been generated that consists of the N-terminal ABD and the first two spectrin repeats of β-spectrin, and the C-terminus of α-spectrin consisting of the last two spectrin repeats and the calmodulin-like domain. This construct was still able to dimerize, bind F-actin, and induce the formation of bundles, but it is unlikely to be functional *in vivo* (Raae *et al.*, 2003). The shock-absorbing role of the spectrin repeats found within these proteins is widely accepted, which leads to the question of why there are so many coiled coils in dystrophin and utrophin. α-Actinin contains only four spectrin repeats, which mediate dimerization and result in a rather inflexible link between the termini of the protein (Ylanne *et al.*, 2001a). In this case, the length of the rod domain clearly defines the distance at which filamentous actin can be crosslinked. α-Actinin is localized to structures that require actin filaments to be crosslinked in either a parallel or antiparallel fashion. This requires the α-actinin dimer to bind actin filaments in orientations separated by as much as 180 degrees. Furthermore, α-actinin is able to accommodate a range of interactin filament crosslinking distances, from 15–40 nm (Liu *et al.*, 2004; Luther, 2000; Taylor *et al.*, 2000). The domain is essential to the formation of the functional dimer and for the separation of the C- and N-termini of the protein, but it would seem that the flexible hinge that separates the rod from the ABD, and the ABD itself, play an important role in determining the ultimate crosslinking distance.

C. *Other Binding Partners*

The repeating constituents of the rod domains of spectrin family proteins were generally regarded as modules for the construction of elongated molecules (Winder, 1997). However, this is not the only function of spectrin repeats. It is widely accepted that proteins containing spectrin repeats are localized to cellular sites that experience significant mechanical stress, and the properties of the spectrin repeat can be used to explain this functionality (see Section IV.B. above). Additionally, some spectrin repeats have acquired functions with a purely structural role; these are able to interact with a variety of structural and signaling proteins (Djinovic-Carugo *et al.*, 2002).

The function of spectrin superfamily proteins is particularly evident when taken in context of their cellular localization. They often form flexible links or structures that allow interactions with the cellular cytoskeletal architecture and the membrane. In both spectrin and dystrophin, such a function is performed, but the spectrin repeats of these molecules are also able to interact with actin and contribute to binding. A portion of the dystrophin rod domain that spans residues 11–17 contains a number of basic repeats that allow a lateral interaction with filamentous actin (Rybakova *et al.*, 2002). The homologous utrophin can also interact laterally with actin. This interaction is distinct from that of dystrophin, as the utrophin rod domain lacks the basic repeat cluster and associates with actin via the first ten spectrin repeats (Rybakova *et al.*, 2002). β-Spectrin also exhibits an extended contact with actin via the first spectrin repeat. In this situation, it was found that the extended contact increased the association of the adjacent ABD with actin (Li and Bennett, 1996). In conjunction with this interaction, it has been found that the second repeat is also required for maximal interaction with adducin (Li and Bennett, 1996), a protein localized at the spectrin-actin junction that is believed to contribute to the assembly of this structure in the membrane skeletal network (Gardner and Bennett, 1987). In the erythrocyte cytoskeletal lattice, β-spectrin interacts with ankyrin, which in turn binds to the cytoplasmic domain of the membrane-associated anion exchanger. This indirect link to the cellular membrane occurs via repeat 15 of β-spectrin (Kennedy *et al.*, 1991) and is largely responsible for the attachment of the spectrin-actin network to the erythrocyte membrane (reviewed in Bennett and Baines, 2001). A much larger number of direct links to transmembrane proteins have been determined for the spectrin repeats of α-actinin (reviewed in Djinovic-Carugo *et al.*, 2002).

The crystal structure of the α-actinin rod domain (Ylanne *et al.*, 2001a) has allowed the analysis of surface features, leading to predictions of

possible protein-protein interaction sites. It was found that the most conserved surface residues were acidic in nature, which would correlate well with the relatively short basic sequences that can be found within the cytoplasmic domains of many transmembrane proteins (Ylanne *et al.*, 2001b). α-Actinin has been found to provide a direct link with a variety of transmembrane proteins including integrins, ICAMs, L-selectin, Ep-Cam, ADAM12, and NMDA receptor subunits (see Djinovic-Carugo *et al.*, 2002 for references). The α-actinin rod domain is also involved in a number of dynamic and regulatory interactions that involve interactions with titin (Young *et al.*, 1998), myotilin (Salmikangas *et al.*, 1999), ALP (Xia *et al.*, 1997), and FATZ (Faulkner *et al.*, 2000) at the Z-disk of striated muscle and interactions with Rho-kinase type protein kinase N (PKN) (Mukai *et al.*, 1997). All of these interactions occur through the rod domain of α-actinin and demonstrate the multivariance of the rod domain as a binding site for the interactions with these proteins (Djinovic-Carugo *et al.*, 2002). Spectrin and dystrophin rod domains have also been demonstrated to interact directly with lipid surfaces, suggesting a lateral association with biological membranes (An *et al.*, 2004; DeWolf *et al.*, 1997; Le Rumeur *et al.*, 2003; Maksymiw *et al.*, 1987).

V. Regulations of Interactions

The spectrin family of proteins, depending on the particular function, has numerous smaller motifs and binding sites for interaction with other proteins. These regions are important, as they are major protein-protein or protein-membrane interaction modules that bind to F-actin, proline-containing ligands, and/or phospholipids. Spectrin and dystrophin/utrophin have all acquired copies of such domains since their evolution from α-actinin, presumably as a consequence of their more diverse roles in the cell.

A. CH Domains

The calponin homology (CH) domain has been identified in many molecules of differing function. However, its presence usually signifies an interaction of some sort with the actin cytoskeleton via an association with F-actin. The domain was initially identified as a 100-residue motif found at the N-terminus of the smooth muscle regulatory protein calponin and, hence, was termed the *CH domain* (Castresana and Saraste, 1995). The refinement of algorithms for the identification of distinct protein motifs has allowed the identification of CH domains in proteins that range

in function from crosslinking to signaling (Korenbaum and Rivero, 2002). Despite the functional variability of this domain, the secondary structure is conserved remarkably well between proteins that contain it (Bramham *et al.*, 2002).

Several mechanisms have been identified that seem to regulate the CH domains found in spectrin, dystrophin, and α-actinin. These range from effects induced by calcium via EF-hand motifs, PIP_2 binding, phosphorylation, and interactions with calmodulin. The actin-binding properties of the nonmuscle isoforms of the F-actin crosslinker α-actinin can be regulated via the presence of EF hands. Calcium does not directly regulate α-actinin's CH domains interaction with F-actin, but it does bind to the EF-hand motif present in the molecule. As α-actinin dimerizes, this brings the CH domains and EF hands in the antiparallel dimer in close association. The conformational changes induced in the EF-hand motif can then exert an effect on the CH domains to influence the interaction with F-actin (Noegel *et al.*, 1987; Tang *et al.*, 2001). α-Actinin has also been found to bind phosphatidylinositol (4,5)-bisphosphate (PIP_2) at the muscle Z-line (Fukami *et al.*, 1992). The PIP_2 binding site has been delineated to a region immediately C-terminal of the third ABS (Fukami *et al.*, 1996), although the precise mechanism of control is not known for this region. It has been found, though, that in nonmuscle cells where α-actinin is associated with actin, this region contained bound PIP_2 whereas free α-actinin did not. This implicates a role of PIP_2 in the activation of α-actinin-induced actin bundling (Fukami *et al.*, 1994).

Calmodulin has also been shown to regulate the interaction of the ABDs from dystrophin (also utrophin) and α-actinin by binding directly to the CH domains (Bonet-Kerrache *et al.*, 1994; Jarrett and Foster, 1995; Winder *et al.*, 1995b) suggesting a potential role for modulating the attachment of these proteins to the cytoskeleton.

Recently, it has been shown that α-actinin is phosphorylated by focal adhesion kinase (FAK) and that this phosphorylation reduces the ability of α-actinin to bind actin (Izaguirre *et al.*, 2001). The site of tyrosine phosphorylation is N-terminal to the first CH domain in a region that is most conserved between spectrin family proteins.

B. EF Hands

EF-hand regions are involved in the chelation of up to two divalent calcium cations (occasionally magnesium) via an interaction through a paired helix-loop-helix structure (Tufty and Kretsinger, 1975). Binding of calcium to this globular domain leads to a dramatic conformational change from ''closed'' to ''open,'' exposing a hydrophobic surface that

binds to a target peptide, often helical in nature. However, divergent evolution has led to a subset of EF hands that no longer chelate calcium and possibly serve an alternate function (Nakayama and Kretsinger, 1994). This is exemplified in α-actinin nonmuscle isoforms, where calcium is bound via the EF hands, thus allowing regulation of actin binding. The muscle-specific isoforms of α-actinin have lost the ability to bind calcium through their EF hands (Blanchard *et al.*, 1989), possibly to protect the muscle architecture from the potential destabilizing effect of calcium during calcium-induced contractions.

Spectrin has also both retained and lost the ability to bind calcium. Calcium and calmodulin bind to human nonerythroid spectrins (αIIβII) at sites that have either degenerated or are absent in erythroid spectrins (αIβI) (Lundberg *et al.*, 1992). The roles of nonerythroid spectrins are far more diverse and, hence, calcium and calmodulin might participate in regulatory events not required in the erythrocyte (Buevich *et al.*, 2004). The EF hands of α-spectrin are brought in close opposition with the β-spectrin ABD once the proteins form the heterodimer. The EF hand is then able to exert regulatory control over the actin-binding activity of the adjacent domain. The molecular details of how this is achieved are still to be determined, however. A similar interface is observed in α-actinin. It is thought that the EF-hand region could engage the actin-binding domain in a manner analogous to calmodulin binding a target peptide. Regulation of the interaction would be affected by the binding of calcium to the EF hand, which would cause a conformational change resulting in altered interaction surfaces. The calcium-binding activity of nonerythroid spectrin has been located to the two EF hands present in the C-terminus of the αII-spectrin (Lundberg *et al.*, 1995; Trave *et al.*, 1995). Buevich and colleagues (2004) found that the EF hands in nonerythroid spectrin exhibited a degree of cooperativity in their binding of calcium, suggesting that EF1 binds before EF2 and modulates the affinity of EF2 for calcium, although overall, calcium binding to α-spectrin has been found to be much weaker than to other EF-hand-containing proteins such as troponin C and calmodulin (Zhang *et al.*, 1995a).

Each of the three EF-hand structures solved from the spectrin family proteins exhibit unique structural and functional differences, even though all are fundamentally similar. The α-spectrin (nonerythroid) EF hands bind calcium and presumably perform some kind of regulatory role regarding the actin-binding function of spectrin (Trave *et al.*, 1995). Due to the low calcium affinity, it is expected that calcium regulatory events involving spectrin would occur in areas of the cell that would experience a transient but significant fluctuation of calcium concentration (Buevich *et al.*, 2004). It is possible that the calcium-bound form of

spectrin in the cell would be stabilized by accessory proteins, as nonerythroid spectrin interacts with many proteins that are involved in regulatory events and not just with the cytoskeleton. In muscle α-actinin (isoform 2), the third and fourth EF hands can be referred to as ''empty'' on the basis of a lack of key liganding residues and large insertions in the helix-loop-helix motif. This muscle isoform of α-actinin is important in striated muscle Z-disk structure, where it interacts with F-actin and titin. The structure of this complex was solved and showed the titin Z-repeat peptide bound in a groove formed by the partially open lobes of the two EF hands (Atkinson *et al.*, 2001). The EF hands of dystrophin were solved as part of a larger structure that also included the adjacent WW domain (Huang *et al.*, 2000). These EF hands had been predicted to be unable to bind calcium due to a lack of key liganding residues (Winder, 1997). The structure of the dystrophin WW-EF region indicated that the EF hands may play a structural role and that they are not required to bind either calcium or a target peptide (Huang *et al.*, 2000). It is still to be elucidated if this region of dystrophin interacts with other target peptides, but as the EF hands are oriented in a closed and compact manner, it is difficult to see how these interactions would occur (Broderick and Winder, 2002). Indeed, studies with constructs spanning both the WW domains and EF-hand regions of dystrophin and utrophin have failed to show any calcium-induced regulation of binding to β-dystroglycan (James *et al.*, 2000; Rentschler *et al.*, 1999).

C. Lipid Binding

Pleckstrin homology (PH) domains are motifs that are approximately 100 amino acids in length and have been identified in over 100 different eukaryotic proteins. They are thought to participate in cell signaling and cytoskeletal regulation via interactions with phospholipids (Lemmon and Ferguson, 1998; reviewed in Rebecchi and Scarlata, 1998). It has been suggested that these domains function as membrane anchors and tethers, as PH domains are often found within membrane-associated proteins (Ferguson *et al.*, 1995). The domain was first recognized in 1993 (Haslam *et al.*, 1993; Mayer *et al.*, 1993; Musacchio *et al.*, 1993) and was quickly followed by the determination of 3-D structures.

The β-spectrin PH domain structure was solved in a lipid-free (Zhang *et al.*, 1995b) and lipid-bound form (Hyvonen *et al.*, 1995). The role of the spectrin PH domain has been proposed to be part of the mechanism whereby spectrin associates directly with the membrane through binding phospholipids. The submembranous framework formed by spectrin is

linked to transmembrane polypeptides via peripheral proteins such as ankyrin and band 4.1 (Viel and Branton, 1996), and spectrin is essential to the integrity of this network. The PH domain of spectrin has been found to have a weak affinity and specificity for $PI(4,5)P_2$ (Harlan *et al.*, 1994; Hyvonen et al., 1995). Lemmon and colleagues (2002) suggested that the β-spectrin/membrane interaction is driven by a delocalized electrostatic attraction between an anionic ligand and the positively charged face of the polarized PH domain. The PH domain of spectrin appears to fall into a class of PH domains that exhibit a moderate affinity for the phosphoinositides. In cells, this polarized domain may direct a few spectrin isoforms to $PI(4,5)P_2$ enriched sites such as caveoli or focal adhesions (Burridge and Chrzanowska-Wodnicka, 1996), where other determinants of membrane association are likely to play an equal or more dominant role in stabilizing attachment. Although membrane attachment is not necessarily dependent on this domain, it has been shown that the PH domain of the human βIΣ2 spectrin isoform binds to protein-depleted membranes containing $PI(4,5)P_2$ and to $Ins(1,4,5)P_3$ in solution (Wang and Shaw, 1995). This domain localizes to plasma membranes in COS7 cells (Wang *et al.*, 1996). $Ins(1,4,5)P_3$ binding has been found to perturb residues located in or near loop 1 of the *Drosophila* spectrin PH domain, as is the case for the N-terminal PH domain and the mouse form of β-spectrin (Zhou *et al.*, 1995). The binding site of β-spectrin has no elaborate hydrogen-bonding network and the inositol ring has no specific contacts with the protein, unlike the PH domain of PLC-δ_1 (Ferguson *et al.*, 1995). Moreover, spectrin does not bind $Ins(1,4,5)P_3$ on the same face as PLC-δ_1, whose binding pocket is located on the other side of the protein.

The β-spectrin PH domain binds weakly to all phosphoinositides and is likely to associate with the negatively charged membrane surface via the positively charged face of the domain. Spectrin networks contain many spectrin molecules, and it is likely that the individual weak association with phosphoinositides is overcome by the overall collective interaction of many molecules. Such a mechanism of multivariant association allows only the assembled cytoskeletal components to interact strongly with cellular membranes, such as in the RBC.

D. Polyproline Binding Domains

The Src homology 3 (SH3) domain of α-spectrin was the first SH3 domain structure to be solved (Musacchio *et al.*, 1992). The domain was initially identified as regions of similar sequence found within signaling proteins, such as the Src family of tyrosine kinases, the Crk adaptor

protein, and phospholipase C-γ (Mayer *et al.*, 1993). In the case of spectrin, the specific target ligand and function are still to be identified. The domain is approximately 60 residues in length and has been identified in many proteins (Bateman *et al.*, 1999; Rubin *et al.*, 2000). The SH3 domain continues to be identified within a variety of proteins; it is subsequently regarded as one of the most common modular protein interaction domains found and is widespread in signaling, adaptor, and cytoskeletal protein alike (reviewed in Mayer, 2001). Due to the small size of this domain, a search for a potential function focused on protein-protein interactions, with screening of expression libraries soon identifying seemingly specific binding partners (Cicchetti *et al.*, 1992). Binding studies have indicated that the interaction sites of SH3 domains were proline-rich, with PxxP being identified as a core conserved binding motif (Ren *et al.*, 1993). It should be noted that profilin and WW domains also make use of a similar mode of interaction with proline-rich helical ligands (Ilsley *et al.*, 2002; Kay *et al.*, 2000; Zarrinpar and Lim, 2000).

As mentioned above, the WW domain is another example of a protein-protein interaction module that binds proline-rich sequences (Kay *et al.*, 2000). Dystrophin and utrophin WW domains interact predominantly with the extracellular matrix receptor dystroglycan, which contains a type 1 WW motif of consensus PPxY (reviewed in Ilsley *et al.*, 2002; Winder, 2001). A structure of a WW domain from dystrophin was solved recently as part of a structure including the EF-hand region, and also with and without a bound β-dystroglycan peptide (Huang *et al.*, 2000).

Chung and Campanelli (1999) found that the interaction between utrophin and β-dystroglycan mirrored that of dystrophin. This is mainly mediated by the WW domain, which recognizes the PPPY peptide at the carboxy terminus of β-dystroglycan. Adhesion-dependent tyrosine phosphorylation of β-dystroglycan within the WW domain binding motif has been found to regulate the WW domain-mediated interaction between utrophin and β-dystroglycan. This was the first demonstration of physiologically relevant tyrosine phosphorylation of a WW domain ligand and parallels the tyrosine phosphorylation of SH3 domain ligands regulating SH3-mediated interactions (James *et al.*, 2000). Investigations performed by Rentschler *et al.* (1999) have also determined that the EF-hand regions following the WW domain are necessary for WW binding. It was later shown that the integrity of the utrophin WW-EF-ZZ region is essential for efficient binding to β-dystroglycan (Tommasi di Vignano *et al.*, 2000). This binding activity can be abolished in utrophin if the ZZ domain is deleted, but only a reduction in binding is observed for dystrophin (Rentschler *et al.*, 1999).

E. ZZ domain

Dystrophin and its autosomal homologue utrophin contain a putative zinc finger motif within their C-terminal cysteine-rich domains, homologous to domains found in sequences of a wide variety of proteins (Ponting *et al.*, 1996). The ZZ domains of dystrophin and utrophin have been shown to bind zinc (Michalak *et al.*, 1996; Winder, 1997) and are believed to be involved in mediating protein-protein interactions, although the precise function of the ZZ domain has not yet been elucidated. It has been found that the cysteine-rich domain of dystrophin is required for binding to β-dystroglycan; it has been shown that the ZZ domain strengthens the interaction between the dystrophin and utrophin WW–EF region with β-dystroglycan (James *et al.*, 2000; Jung *et al.*, 1995; Rentschler *et al.*, 1999). More recently, Ishikawa-Sakurai and colleagues identified the components of the C-terminal domain of dystrophin that are required for the full binding activity. They have detailed the extent of the C-terminal sequence (residues 3026–3345) that is required for effective binding and have identified cysteine 3340 within the ZZ domain as essential to the binding activity with β-dystroglycan (Ishikawa-Sakurai *et al.*, 2004). The functional importance of the ZZ domain has been proven further by the identification of a rare mutation where C3340 has been mutated to a tyrosine, resulting in the affected individual suffering from a form of DMD (Lenk *et al.* 1996). However, to date, no structure of any ZZ domain has been solved. C-terminal to the ZZ domain is a pair of highly conserved helices predicted to form dimeric coiled-coils (Blake *et al.*, 1995). These helices, which are restricted to dystrophin family proteins (dystrophin, utrophin, DRP2, and dystrobrevin), are involved in heterophilic-associations between family members and also members of the syntrophin family of proteins (reviewed in Blake *et al.*, 2002).

VI. Disease

A. Duchenne Muscular Dystrophy

DMD is a severe X-linked recessive, progressive muscle wasting disease that affects approximately 1 in 3500 newborn males (Emery, 1991). An allelic variant of DMD is also known, referred to as *Becker muscular dystrophy* (BMD). It has a later onset and lesser phenotype than DMD, resulting in longer life expectancy (reviewed in O'Brien and Kunkel, 2001). DMD is caused by mutations in the DMD gene that encodes the cytoskeletal linker protein dystrophin. The vast majority of DMD mutations result in the

complete absence of dystrophin, whereas a truncated protein is often associated with the milder Becker form of the disease (Kingston *et al.*, 1983). Mutations in the genes encoding other components of the dystrophin-associated protein complex cause other forms of dystrophy, such as limb-girdle and congenital dystrophies.

The cause of approximately 65% of DMD pathologies can be traced to large deletions or duplications within the dystrophin gene. The remaining cases are the result of small insertion/deletion mutations and point mutations (Koenig *et al.*, 1989; Monaco *et al.*, 1985; reviewed by Roberts *et al.*, 1994). In DMD, it has been found that point mutations nearly always result in a truncation of the open reading frame causing nonsense-mediated decay, but rare cases are known where a truncated nonfunctional protein is transcribed (Kerr *et al.*, 2001). In BMD, most point mutations disrupt splicing, which results in an intact but interstitially deleted open reading frame and a partially functional protein (reviewed in Roberts *et al.*, 1994).

Mutations identified in all major domains of dystrophin result in disease phenotypes ranging from mild to severe. Beginning with the N-terminus, mutations have been identified that stem from missense and inframe mutations. In the ABD, a missense mutation resulting in an amino acid change of an arginine residue for leucine 54 results in a DMD phenotype with reduced levels of protein (Prior *et al.*, 1995). DMD patients have also been described with inframe deletions of exons 3–25, indirectly resulting in normal levels of truncated protein (Vainzof *et al.*, 1993).

The rod domain of dystrophin has been found to accommodate large inframe deletions. A case where a patient was found to be missing exons 17–48, corresponding to a 73% deletion of the rod domain, only exhibited a mild form of BMD (England *et al.*, 1990). Large deletions of the rod domain have also been observed in other BMD patients (Love *et al.*, 1991; Winnard *et al.*, 1993), the phenotypes of which are usually milder than those of DMD.

Few missense mutations have been described in DMD patients, although two informative substitutions have been identified in the cysteine-rich domain. The cysteine-rich domain contains a number of motifs that are important for regulation and protein-protein interactions. The substitution of a conserved cysteine residue for a tyrosine at position 3340 results in reduced but detectable levels of dystrophin. This mutation alters one of the coordinating residues in the ZZ domain that is thought to interfere with the binding of the dystrophin-associated protein β-dystroglycan (Lenk *et al.*, 1996). Another substitution involving an aspartate to a histidine at position 3335 is also thought to affect the β-dystroglycan binding site (Goldberg *et al.*, 1998). Removal of a highly conserved

glutamic acid (residue 3367) adjacent to the dystrophin ZZ domain results in a phenotype of DMD with substantial retention of a presumably functionally compromised dystrophin protein (Becker *et al.*, 2003). Interestingly, the cysteine-rich domain is never deleted in BMD patients, suggesting that this domain is critical for dystrophin function as the BMD phenotypes are less severe (Rafael *et al.*, 1996).

A small number of cases have been identified in which there is a deletion of the carboxy-terminus of dystrophin. In these patients, it is common for the mutant protein to localize to the sarcolemma (Bies *et al.*, 1992; Helliwell *et al.*, 1992; Hoffman *et al.*, 1991). These cases are good examples of the importance of the cysteine-rich and C-terminal domains of dystrophin, presumably reflecting the importance of interactions with components of the dystrophin-associated glycoprotein complex. Many single point mutations within dystrophin are also known.

B. Hereditary Spherocytosis

The membrane skeleton acts as an elastic semisolid, allowing brief periods of deformation followed by reestablishment of the original cell shape (reviewed by Bennett and Gilligan, 1993). Erythrocytes in the human bloodstream have to squeeze repeatedly through narrow capillaries of diameters smaller than their own dimensions while resisting rupture. A functional erythrocyte membrane is pivotal to maintaining the functional properties of the erythrocyte. This importance is apparent when examination is made of many hemolytic anemias, where mutation of proteins involved in the structure of the submembranous cytoskeleton, and its attachment to the lipid bilayer, result in a malformed or altered cytoskeletal architecture and a disease phenotype.

Hereditary spherocytosis (HS) comprises a group of inherited hemolytic anemias characterized by chronic hemolysis with a broad spectrum of severity (Hassoun *et al.*, 1997). The principal cellular defect is the loss of erythrocyte surface area relative to the intracellular volume, although increased osmotic frailty is also a factor. A distinctive spherical red blood cell (RBC) morphology is observed in sufferers of HS and splenic destruction of these abnormal erythrocytes is the primary cause of the hemolysis experienced (Delaunay, 1995; Palek and Jarolim, 1993).

The primary biochemical defects of HS are linked to proteins important to the interaction between the membrane skeleton and the lipid bilayer involving α- and β-spectrin, ankyrin, band 3, and protein 4.2 (Gallagher and Forget, 1998). Combined spectrin and ankyrin deficiency (Coetzer *et al.*, 1988; Pekrun *et al.*, 1993; Savvides *et al.*, 1993) is most commonly observed, followed by band 3 deficiency (Iolascon *et al.*, 1992; Jarolim *et al.*,

1990; Lux *et al.*, 1990), isolated spectrin deficiency (Agre *et al.*, 1986; Eber *et al.*, 1990), and then protein 4.2 deficiency (Bouhassira *et al.*, 1992; Ghanem *et al.*, 1990; Rybicki *et al.*, 1988).

Spectrin forms an integral part of the erythrocyte cytoskeletal architecture; any defects that disrupt the association of the spectrin heterotetramer or the interaction with any of the other submembranous proteins can result in RBC defects (reviewed in Hassoun and Palek, 1996). Indeed, abnormalities of the β-spectrin N-terminus and the α-spectrin C-terminus affect the self-association site and result in hereditary elliptocytosis and hereditary pyropoikilocytosis (Delaunay, 1995; Delaunay and Dhermy, 1993; Palek and Jarolim, 1993).

Defects outside the self-association site of spectrin are also associated with HS. In many sufferers of HS, both dominant and recessive forms of the disease result in spectrin deficiency. Normally α-spectrin is synthesized in a three- or fourfold excess of β-spectrin (Hanspal and Palek, 1987, 1992). Heterozygotes for α-spectrin should still be able to produce enough normal α-spectrin to associate with the majority of β-spectrin present in the RBC. The deficiency would only become apparent in sufferers who are homozygotes or compound heterozygotes for α-spectrin gene mutations where there would be insufficient α-spectrin to associate with β-spectrin. In the case of β-spectrin deficiency, mutations of the β-spectrin gene are associated with dominantly inherited HS because β-spectrin is the limiting component in heterotetramer formation (Gallagher and Forget, 1998). Redundant unassembled α-spectrin chains are degraded in the lysosomal compartment (Woods and Lazarides, 1985). As a result of this, the synthesis of β-spectrin would seem to be the rate-limiting factor for the assembly of the spectrin $\alpha\beta$ heterodimers on the membrane. Therefore, in contrast to α-spectrin, defects of β-spectrin are more likely to be expressed in the heterozygous state (Hassoun *et al.*, 1996). The rate of turnover of the β-spectrin subunit is viewed as critical to the posttranscriptional regulation of the heterodimer assembly on the membrane (Moon and Lazarides, 1983; Woods and Lazarides, 1985).

The most common mutations associated with HS are those involving ankyrin. However, many mutations have been identified that involve spectrin and result in either splicing, missense, deletion, and insertion mutations (see references within Gallagher and Forget, 1998). Of these mutations, two were found in the second CH domain, T182G and I220V, with both residues being important for maintaining the hydrophobic core of the CH domain (Becker *et al.*, 1993; Hassoun *et al.*, 1997). Alteration of Trp 182 to Gly in the short helix B, in particular, would be expected to have a severe effect on the stability of the CH domain with consequent effects on the function of the whole ABD. Another isolated

β-spectrin point mutation is also known and results in defective binding to protein 4.1 (Becker *et al.*, 1993; Wolfe *et al.*, 1982). Further investigation suggested that the mutant spectrin was likely to be unstable and susceptible to proteolytic degradation, leading to an overall spectrin deficiency (Becker *et al.*, 1993).

Mutations within the rod domain of β-spectrin have also been identified. A truncated β-spectrin lacking repeat 12 and part of repeat 13 is the product of a large genomic deletion (Hassoun *et al.*, 1995). Studies of the mutated spectrin have shown synthesis and stability were unaffected by the deletion, but its incorporation into the membrane skeleton was hampered in erythroblasts. The misincorporation is likely to be a result of defective binding to ankyrin (Hassoun *et al.*, 1995) as the missing section of the rod domain is in close proximity to the ankyrin-binding domain (Kennedy *et al.*, 1991). A second mutation of β-spectrin characterized by truncation and overall deficiency was found to involve a point mutation within intron 17 (Hassoun *et al.*, 1996). This mutation leads to an unstable transcriptional message that lacks exons 16 and 17. It was proposed that the mutant might be susceptible to proteolytic degradation in the cytoplasm of erythroblasts, ultimately contributing to a lack of spectrin at the membrane. Several other frameshift, missense, and nonsense mutations of the β-spectrin gene have also been discovered in patients suffering from HS (Hassoun *et al.*, 1995).

C. *Familial Focal Segmental Glomerulosclerosis*

In humans, mutations in the ACTN4 gene located on chromosome 19q13 result in α-actinin-4 mutations, which are believed to cause a form of familial focal segmental glomerulosclerosis (FSGS) (Kaplan *et al.*, 2000). FSGS is a common nonspecific renal lesion characterized by regions of sclerosis in some renal glomeruli, often resulting in loss of kidney function and ultimately end-stage renal failure. FSGS is often secondary to other disorders such as HIV infection, obesity, hypertension, and diabetes, but can also appear as an isolated idiopathic condition. When FSGS occurs as a primary process, it is thought to result from a defect in glomerular podocyte function (Ichikawa and Fogo, 1996). α-Actinin 4 has been implicated in some cases of autosomal dominant FSGS due to the high expression level of this protein in the glomerular podocyte and also its early upregulation during the course of nephritic syndrome in some animal models (Drenckhahn and Franke, 1988; Shirato *et al.*, 1996; Smoyer *et al.*, 1997).

Kaplan and colleagues (2000) have sequenced the coding region of ACTN4 from a number of families that present one form or another of

FSGS. Three specific residues were characterized—K228E, T232I, and S235P—and can be found on the solvent-accessible surface in helix G of the second CH domain (Kaplan *et al.*, 2000). These mutations are not expected to perturb the secondary structure of the actin-binding domain. However, they are all highly conserved among all four human isoforms of α-actinin, as well as α-actinins of other species. These mutations are not in a region of CH2 that is thought to directly interact with F-actin. However, their presence has a marked effect on the functionality of the mutant protein when in the cellular environment and when binding to F-actin. Actin-binding experiments using the mutant α-actinin have revealed that more of the mutant protein associates with F-actin than when the wild type is used. This observation leads to the conclusion that the ACTN4 mutations are able to cause or contribute to the susceptibility to focal segmental glomerulosclerosis (Kaplan *et al.*, 2000). The increased affinity of α-actinin-4 for F-actin was also confirmed by Michaud *et al.* (2003). Further investigations performed by Yao and colleagues (2004) have demonstrated that the mutant α-actinins exhibit alter structural characteristics, localize abnormally, and are targets for degradation. They suggest that the mutant α-actinin-4 is much less dynamic within the cellular environment and, due to its propensity to aggregate, loss of normal function becomes inevitable and contributes to progression of kidney disease. It is possible that the effects of these mutations are more apparent in the kidney due to the high level of expression of α-actinin 4 in the podocyte (Drenckhahn *et al.*, 1990; Kurihara *et al.*, 1995). These dominant mutations may also alter the manner in which α-actinin interacts with other cytoskeletal components in conjunction with the effects on the normal assembly and disassembly of actin.

VII. Conclusion

The spectrin superfamily is a group of cytoskeletal proteins that have been found to perform a variety of cellular functions. The role of each protein and their interactions within the cellular environment stem from the specific domains found within each protein and the manner in which they are organized. Each of the family members is formed from discrete modular domains that have the ability to interact or modulate specific interactions or impart physical abilities on the protein relevant to its function. The particular members of this protein family have been shown to be evolutionary related. α-Actinin is believed to be the ancestor of the whole group and, indeed, sequence and phylogenetic analysis has found this to be the case. It is astounding that from a simple precursor containing few domains such a family of functionally diverse proteins can be

generated. It should also be noted that the domains present within the proteins of the spectrin superfamily have been adopted by many other protein groups, further supporting their suitability to perform the interactions that they have evolved. There is still much more to be learned from this group of proteins, including not only their structure-function relationships, but also the manner in which they fit within the cellular environment as a whole.

References

Agre, P., Casella, J. F., Zinkham, W. H., McMillan, C., and Bennett, V. (1985). Partial deficiency of erythrocyte spectrin in hereditary spherocytosis. *Nature* **314,** 380–383.

Agre, P., Asimos, A., Casella, J. F., and McMillan, C. (1986). Inheritance pattern and clinical response to splenectomy as a reflection of erythrocyte spectrin deficiency in hereditary spherocytosis. *N. Engl. J. Med.* **315,** 1579–1583.

Amann, K. J., Guo, A. W., and Ervasti, J. M. (1999). Utrophin lacks the rod domain actin binding activity of dystrophin. *J. Biol. Chem.* **274,** 35375–35380.

An, X., Guo, X., Sum, H., Morrow, J., Gratzer, W., and Mohandas, N. (2004). Phosphatidylserine binding sites in erythroid spectrin: Location and implications for membrane stability. *Biochemistry* **43,** 310–315.

Atkinson, R. A., Joseph, C., Kelly, G., Muskett, F. W., Frenkiel, T. A., Nietlispach, D., and Pastore, A. (2001). Ca2+-independent binding of an EF-hand domain to a novel motif in the alpha-actinin-titin complex. *Nat. Struct. Biol.* **8,** 853–857.

Baines, A. J. (2003). Comprehensive analysis of all triple helical repeats in beta-spectrins reveals patterns of selective evolutionary conservation. *Cell Mol. Biol. Lett.* **8,** 195–214.

Banuelos, S., Saraste, M., and Carugo, K. D. (1998). Structural comparisons of calponin homology domains: Implications for actin binding. *Structure* **6,** 1419–1431.

Baron, M. D., Davison, M. D., Jones, P., Patel, B., and Critchley, D. R. (1987). Isolation and characterization of a cDNA encoding a chick alpha-actinin. *J. Biol. Chem.* **262,** 2558–2561.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D., and Sonnhammer, E. L. (1999). Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.* **27,** 260–262.

Becker, P. S., Tse, W. T., Lux, S. E., and Forget, B. G. (1993). Beta spectrin kissimmee: A spectrin variant associated with autosomal dominant hereditary spherocytosis and defective binding to protein 4.1. *J. Clin. Invest.* **92,** 612–616.

Becker, K., Robb, S. A., Hatton, Z., Yau, S. C., Abbs, S., and Roberts, R. G. (2003). Loss of a single amino acid from dystrophin resulting in Duchenne muscular dystrophy with retention of dystrophin protein. *Hum. Mutat.* **21,** 651.

Bennett, H., and Condeelis, J. (1988). Isolation of an immunoreactive analogue of brain fodrin that is associated with the cell cortex of Dictyostelium amoebae. *Cell Motil. Cytoskeleton* **11,** 303–317.

Bennett, V., and Gilligan, D. M. (1993). The spectrin-based membrane skeleton and micron-scale organization of the plasma membrane. *Annu. Rev. Cell Biol.* **9,** 27–66.

Bennett, V., and Baines, A. J. (2001). Spectrin and ankyrin-based pathways: Metazoan inventions for integrating cells into tissues. *Physiol. Rev.* **81,** 1353–1392.

Bies, R. D., Caskey, C. T., and Fenwick, R. (1992). An intact cysteine-rich domain is required for dystrophin function. *J. Clin. Invest.* **90,** 666–672.

Blake, D. J., Tinsley, J. M., Davies, K. E., Knight, A. E., Winder, S. J., and Kendrick-Jones, J. (1995). Coiled-coil regions in the carboxy-terminal domains of dystrophin and related proteins: Potentials for protein-protein interactions. *Trends Biochem. Sci.* **20,** 133–135.

Blake, D. J., Weir, A., Newey, S. E., and Davies, K. E. (2002). Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiol. Rev.* **82,** 291–329.

Blanchard, A., Ohanian, V., and Critchley, D. (1989). The structure and function of alpha-actinin. *J. Muscle Res. Cell Motil.* **10,** 280–289.

Bonet-Kerrache, A., Fabbrizio, E., and Mornet, D. (1994). N-terminal domain of dystrophin. *FEBS Lett.* **355,** 49–53.

Bouhassira, E. E., Schwartz, R. S., Yawata, Y., Ata, K., Kanzaki, A., Qiu, J. J., Nagel, R. L., and Rybicki, A. C. (1992). An alanine-to-threonine substitution in protein 4.2 cDNA is associated with a Japanese form of hereditary hemolytic anemia (protein 4.2NIPPON). *Blood* **79,** 1846–1854.

Bramham, J., Hodgkinson, J. L., Smith, B. O., Uhrin, D., Barlow, P. N., and Winder, S. J. (2002). Solution structure of the calponin CH domain and fitting to the 3D-helical reconstruction of F-actin:calponin. *Structure* **10,** 249–258.

Bresnick, A. R., Warren, V., and Condeelis, J. (1990). Identification of a short sequence essential for actin binding by Dictyostelium ABP-120. *J. Biol. Chem.* **265,** 9236–9240.

Bresnick, A. R., Janmey, P. A., and Condeelis, J. (1991). Evidence that a 27-residue sequence is the actin-binding site of ABP-120. *J. Biol. Chem.* **266,** 12989–12993.

Broderick, M. J. F., and Winder, S. J. (2002). Towards a complete atomic structure of spectrin family proteins. *J. Struct. Biol.* **137,** 184–193.

Buevich, A. V., Lundberg, S., Sethson, I., Edlund, U., and Backman, L. (2004). NMR studies of calcium-binding to mutant alpha-spectrin EF-hands. *Cell Mol. Biol. Lett.* **9,** 167–186.

Burridge, K., and Chrzanowska-Wodnicka, M. (1996). Focal adhesions, contractility, and signaling. *Annu. Rev. Cell Dev. Biol.* **12,** 463–518.

Byers, T. J., Husain-Chishti, A., Dubreuil, R. R., Branton, D., and Goldstein, L. S. (1989). Sequence similarity of the amino-terminal domain of *Drosophila* beta spectrin to alpha actinin and dystrophin. *J. Cell Biol.* **109,** 1633–1641.

Byers, T. J., Brandin, E., Lue, R. A., Winograd, E., and Branton, D. (1992). The complete sequence of *Drosophila* beta-spectrin reveals supra-motifs comprising eight 106-residue segments. *Proc. Natl. Acad. Sci. USA* **89,** 6187–6191.

Campbell, K. P., and Kahl, S. D. (1989). Association of dystrophin and an integral membrane glycoprotein. *Nature* **338,** 259–262.

Carugo, K. D., Banuelos, S., and Saraste, M. (1997). Crystal structure of a calponin homology domain. *Nat. Struct. Biol.* **4,** 175–179.

Castresana, J., and Saraste, M. (1995). Does Vav bind to F-actin through a CH domain? *FEBS Lett.* **374,** 149–151.

Chung, W., and Campanelli, J. T. (1999). WW and EF hand domains of dystrophin-family proteins mediate dystroglycan binding. *Mol. Cell Biol. Res. Commun.* **2,** 162–171.

Cicchetti, P., Mayer, B. J., Thiel, G., and Baltimore, D. (1992). Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science* **257,** 803–806.

Claudepierre, T., Rodius, F., Frasson, M., Fontaine, V., Picaud, S., Dreyfus, H., Mornet, D., and Rendon, A. (1999). Differential distribution of dystrophins in rat retina. *Invest. Ophthalmol. Vis. Sci.* **40,** 1520–1529.

Coetzer, T. L., Lawler, J., Liu, S. C., Prchal, J. T., Gualtieri, R. J., Brain, M. C., Dacie, J. V., and Palek, J. (1988). Partial ankyrin and spectrin deficiency in severe, atypical hereditary spherocytosis. *N. Engl. J. Med.* **318,** 230–234.

Coffey, A. J., Roberts, R. G., Green, E. D., Cole, C. G., Butler, R., Anand, R., Giannelli, F., and Bentley, D. R. (1992). Construction of a 2.6-Mb contig in yeast artificial chromosomes spanning the human dystrophin gene using an STS-based approach. *Genomics* **12,** 474–484.

Craig, S. W., and Pardo, J. V. (1983). Gamma actin, spectrin, and intermediate filament proteins colocalize with vinculin at costameres, myofibril-to-sarcolemma attachment sites. *Cell Motil.* **3,** 449–462.

Critchley, D. R. (2000). Focal adhesions – The cytoskeletal connection. *Curr. Opin. Cell Biol.* **12,** 133–139.

Danowski, B. A., Imanaka-Yoshida, K., Sanger, J. M., and Sanger, J. W. (1992). Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes. *J. Cell Biol.* **118,** 1411–1420.

Davison, M. D., and Critchley, D. R. (1988). Alpha-actinins and the DMD protein contain spectrin-like repeats. *Cell* **52,** 159–160.

Delaunay, J. (1995). Genetic disorders of the red cell membranes. *FEBS Lett.* **369,** 34–37.

Delaunay, J., and Dhermy, D. (1993). Mutations involving the spectrin heterodimer contact site: Clinical expression and alterations in specific function. *Semin. Hematol.* **30,** 21–33.

DeWolf, C., McCauley, P., Sikorski, A. F., Winlove, C. P., Bailey, A. I., Kahana, E., Pinder, J. C., and Gratzer, W. B. (1997). Interaction of dystrophin fragments with model membranes. *Biophys. J.* **72,** 2599–2604.

Djinovic-Carugo, K., Gautel, M., Ylanne, J., and Young, P. (2002). The spectrin repeat: A structural platform for cytoskeletal protein assemblies. *FEBS Lett.* **513,** 119–123.

Drenckhahn, D., and Franke, R. P. (1988). Ultrastructural organization of contractile and cytoskeletal proteins in glomerular podocytes of chicken, rat, and man. *Lab. Invest.* **59,** 673–682.

Drenckhahn, D., Schnittler, H., Nobiling, R., and Kriz, W. (1990). Ultrastructural organization of contractile proteins in rat glomerular mesangial cells. *Am. J. Pathol.* **137,** 1343–1351.

Dubreuil, R. R. (1991). Structure and evolution of the actin crosslinking proteins. *Bioessays* **13,** 219–226.

Dubreuil, R. R., Byers, T. J., Sillman, A. L., Bar-Zvi, D., Goldstein, L. S., and Branton, D. (1989). The complete sequence of *Drosophila* alpha-spectrin: Conservation of structural domains between alpha-spectrins and alpha-actinin. *J. Cell Biol.* **109,** 2197–2205.

Earnest, J. P., Santos, G. F., Zuerbig, S., and Fox, J. E. (1995). Dystrophin-related protein in the platelet membrane skeleton. Integrin-induced change in detergent-insolubility and cleavage by calpain in aggregating platelets. *J. Biol. Chem.* **270,** 27259–27265.

Eber, S. W., Armbrust, R., and Schroter, W. (1990). Variable clinical severity of hereditary spherocytosis: Relation to erythrocytic spectrin concentration, osmotic fragility, and autohemolysis. *J. Pediatr.* **117,** 409–416.

Ehmer, S., Herrmann, R., Bittner, R., and Voit, T. (1997). Spatial distribution of beta-spectrin in normal and dystrophic human skeletal muscle. *Acta. Neuropathol. (Berl.)* **94,** 240–246.

Emery, A. E. (1991). Population frequencies of inherited neuromuscular diseases – A world survey. *Neuromuscul. Disord.* **1,** 19–29.

England, S. B., Nicholson, L. V., Johnson, M. A., Forrest, S. M., Love, D. R., Zubrzycka-Gaarn, E. E., Bulman, D. E., Harris, J. B., and Davies, K. E. (1990). Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* **343,** 180–182.

Ervasti, J. M., and Campbell, K. P. (1993a). Dystrophin and the membrane skeleton. *Curr. Opin. Cell Biol.* **5,** 82–87.

Ervasti, J. M., and Campbell, K. P. (1993b). A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J. Cell Biol.* **122,** 809–823.

Evans, E. A., and Hochmuth, R. M. (1977). A solid-liquid composite model of the red cell membrane. *J. Membr. Biol.* **30,** 351–362.

Faulkner, G., Pallavicini, A., Comelli, A., Salamon, M., Bortoletto, G., Ievolella, C., Trevisan, S., Kojic, S., Dalla Vecchia, F., Laveder, P., Valle, G., and Lanfranchi, G. (2000). FATZ, a filamin-, actinin-, and telethonin-binding protein of the Z-disc of skeletal muscle. *J. Biol. Chem.* **275,** 41234–41242.

Ferguson, K. M., Lemmon, M. A., Sigler, P. B., and Schlessinger, J. (1995). Scratching the surface with the PH domain. *Nat. Struct. Biol.* **2,** 715–718.

Fukami, K., Furuhashi, K., Inagaki, M., Endo, T., Hatano, S., and Takenawa, T. (1992). Requirement of phosphatidylinositol 4,5-bisphosphate for alpha-actinin function. *Nature* **359,** 150–152.

Fukami, K., Endo, T., Imamura, M., and Takenawa, T. (1994). Alpha-actinin and vinculin are PIP2-binding proteins involved in signaling by tyrosine kinase. *J. Biol. Chem.* **269,** 1518–1522.

Fukami, K., Sawada, N., Endo, T., and Takenawa, T. (1996). Identification of a phosphatidylinositol 4,5-bisphosphate-binding site in chicken skeletal muscle alpha-actinin. *J. Biol. Chem.* **271,** 2646–2650.

Galkin, V. E., Orlova, A., VanLoock, M. S., Rybakova, I. N., Ervasti, J. M., and Egelman, E. H. (2002). The utrophin actin-binding domain binds F-actin in two different modes: Implications for the spectrin superfamily of proteins. *J. Cell Biol.* **157,** 243–251.

Galkin, V. E., Orlova, A., VanLoock, M. S., and Egelman, E. H. (2003). Do the utrophin tandem calponin homology domains bind F-actin in a compact or extended conformation? *J. Mol. Biol.* **331,** 967–972.

Gallagher, P. G., and Forget, B. G. (1998). Hematologically important mutations: Spectrin and ankyrin variants in hereditary spherocytosis. *Blood Cells Mol. Dis.* **24,** 539–543.

Garcia-Alvarez, B., Bobkov, A., Sonnenberg, A., and de Pereda, J. M. (2003). Structural and functional analysis of the actin binding domain of plectin suggests alternative mechanisms for binding to F-actin and integrin beta4. *Structure* **11,** 615–625.

Gardner, K., and Bennett, V. (1987). Modulation of spectrin-actin assembly by erythrocyte adducin. *Nature* **328,** 359–362.

Ghanem, A., Pothier, B., Marechal, J., Ducluzeau, M. T., Morle, L., Alloisio, N., Feo, C., Ben Abdeladhim, A., Fattoum, S., and Delaunay, J. (1990). A haemolytic syndrome associated with the complete absence of red cell membrane protein 4.2 in two Tunisian siblings. *Br. J. Haematol.* **75,** 414–420.

Gimona, M., and Mital, R. (1998). The single CH domain of calponin is neither sufficient nor necessary for F-actin binding. *J. Cell Sci.* **111,** 1813–1821.

Gimona, M., and Winder, S. J. (1998). Single calponin homology domains are not actin-binding domains. *Curr. Biol.* **8,** R674–675.

Goldberg, L. R., Hausmanowa-Petrusewicz, I., Fidzianska, A., Duggan, D. J., Steinberg, L. S., and Hoffman, E. P. (1998). A dystrophin missense mutation showing persistence of dystrophin and dystrophin-associated proteins yet a severe phenotype. *Ann. Neurol.* **44,** 971–976.

Goldsmith, S. C., Pokala, N., Shen, W., Fedorov, A. A., Matsudaira, P., and Almo, S. C. (1997). The structure of an actin-crosslinking domain from human fimbrin. *Nat. Struct. Biol.* **4,** 708–712.

Gratzer, W. B. (1985). Preparation of spectrin. *Methods Enzymol.* **85**(Pt. B), 475–480.

Grum, V. L., Li, D., MacDonald, R. I., and Mondragon, A. (1999). Structures of two repeats of spectrin suggest models of flexibility. *Cell* **98,** 523–535.

Hanein, D., Volkmann, N., Goldsmith, S., Michon, A. M., Lehman, W., Craig, R., DeRosier, D., Almo, S., and Matsudaira, P. (1998). An atomic model of fimbrin binding to F-actin and its implications for filament crosslinking and regulation. *Nat. Struct. Biol.* **5,** 787–792.

Hanspal, M., and Palek, J. (1987). Synthesis and assembly of membrane skeletal proteins in mammalian red cell precursors. *J. Cell Biol.* **105,** 1417–1424.

Hanspal, M., and Palek, J. (1992). Biogenesis of normal and abnormal red blood cell membrane skeleton. *Semin. Hematol.* **29,** 305–319.

Harlan, J. E., Hajduk, P. J., Yoon, H. S., and Fesik, S. W. (1994). Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature* **371,** 168–170.

Haslam, R. J., Koide, H. B., and Hemmings, B. A. (1993). Pleckstrin domain homology. *Nature* **363,** 309–310.

Hassoun, H., and Palek, J. (1996). Hereditary spherocytosis: A review of the clinical and molecular aspects of the disease. *Blood Rev.* **10,** 129–147.

Hassoun, H., Vassiliadis, J. N., Murray, J., Yi, S. J., Hanspal, M., Johnson, C. A., and Palek, J. (1996). Hereditary spherocytosis with spectrin deficiency due to an unstable truncated beta spectrin. *Blood* **87,** 2538–2545.

Hassoun, H., Vassiliadis, J. N., Murray, J., Yi, S. J., Hanspal, M., Ware, R. E., Winter, S. S., Chiou, S. S., and Palek, J. (1995). Molecular basis of spectrin deficiency in beta spectrin Durham. A deletion within beta spectrin adjacent to the ankyrin-binding site precludes spectrin attachment to the membrane in hereditary spherocytosis. *J. Clin. Invest.* **96,** 2623–2629.

Hassoun, H., Vassiliadis, J. N., Murray, J., Njolstad, P. R., Rogus, J. J., Ballas, S. K., Schaffer, F., Jarolim, P., Brabec, V., and Palek, J. (1997). Characterization of the underlying molecular defect in hereditary spherocytosis associated with spectrin deficiency. *Blood* **90,** 398–406.

Helliwell, T. R., Ellis, J. M., Mountford, R. C., Appleton, R. E., and Morris, G. E. (1992). A truncated dystrophin lacking the C-terminal domains is localized at the muscle membrane. *Am. J. Hum. Genet.* **50,** 508–514.

Hoffman, E. P., Garcia, C. A., Chamberlain, J. S., Angelini, C., Lupski, J. R., and Fenwick, R. (1991). Is the carboxyl-terminus of dystrophin required for membrane association? A novel, severe case of Duchenne muscular dystrophy. *Ann. Neurol.* **30,** 605–610.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* **7,** 634–638.

Hyvonen, M., Macias, M. J., Nilges, M., Oschkinat, H., Saraste, M., and Wilmanns, M. (1995). Structure of the binding site for inositol phosphates in a PH domain. *EMBO J.* **14,** 4676–4685.

Ichikawa, I., and Fogo, A. (1996). Focal segmental glomerulosclerosis. *Pediatr. Nephrol.* **10,** 374–391.

Ilsley, J. L., Sudol, M., and Winder, S. J. (2002). The WW domain: Linking cell signaling to the membrane cytoskeleton. *Cell Signal.* **14,** 183–189.

Iolascon, A., Miraglia del Giudice, E., Perrotta, S., Pinto, L., Fiorelli, G., Cappellini, D. M., Vasseur, C., Bursaux, E., and Cutillo, S. (1992). Hereditary spherocytosis (HS) due to loss of anion exchange transporter. *Haematologica* **77,** 450–456.

Ishikawa-Sakurai, M., Yoshida, M., Imamura, M., Davies, K. E., and Ozawa, E. (2004). ZZ domain is essentially required for the physiological binding of dystrophin and utrophin to beta-dystroglycan. *Hum. Mol. Genet.* **13,** 693–702.

Izaguirre, G., Aguirre, L., Hu, Y. P., Lee, H. Y., Schlaepfer, D. D., Aneskievich, B. J., and Haimovich, B. (2001). The cytoskeletal/non-muscle isoform of alpha-actinin is phosphorylated on its actin-binding domain by the focal adhesion kinase. *J. Biol. Chem.* **276,** 28676–28685.

James, M., Nuttall, A., Ilsley, J. L., Ottersbach, K., Tinsley, J. M., Sudol, M., and Winder, S. J. (2000). Adhesion-dependent tyrosine phosphorylation of (beta)-dystroglycan regulates its interaction with utrophin. *J. Cell Sci.* **113,** 1717–1726.

Jarolim, P., Lahav, M., Liu, S. C., and Palek, J. (1990). Effect of hemoglobin oxidation products on the stability of red cell membrane skeletons and the associations of skeletal proteins: Correlation with a release of hemin. *Blood* **76,** 2125–2131.

Jarrett, H. W., and Foster, J. L. (1995). Alternate binding of actin and calmodulin to multiple sites on dystrophin. *J. Biol. Chem.* **270,** 5578–5586.

Jung, D., Yang, B., Meyer, J., Chamberlain, J. S., and Campbell, K. P. (1995). Identification and characterization of the dystrophin anchoring site on beta-dystroglycan. *J. Biol. Chem.* **270,** 27305–27310.

Kaplan, J. M., Kim, S. H., North, K. N., Rennke, H., Correia, L. A., Tong, H. Q., Mathis, B. J., Rodriguez-Perez, J. C., Allen, P. G., Beggs, A. H., and Pollak, M. R. (2000). Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis. *Nat. Genet.* **24,** 251–256.

Kay, B. K., Williamson, M. P., and Sudol, M. (2000). The importance of being proline: The interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14,** 231–241.

Keep, N. H., Norwood, F. L., Moores, C. A., Winder, S. J., and Kendrick-Jones, J. (1999a). The 2.0Å structure of the second calponin homology domain from the actin-binding region of the dystrophin homologue utrophin. *J. Mol. Biol.* **285,** 1257–1264.

Keep, N. H., Winder, S. J., Moores, C. A., Walke, S., Norwood, F. L., and Kendrick-Jones, J. (1999b). Crystal structure of the actin-binding region of utrophin reveals a head-to-tail dimer. *Structure Fold. Des.* **7,** 1539–1546.

Kennedy, S. P., Warren, S. L., Forget, B. G., and Morrow, J. S. (1991). Ankyrin binds to the 15th repetitive unit of erythroid and nonerythroid beta-spectrin. *J. Cell Biol.* **115,** 267–277.

Kerr, T. P., Sewry, C. A., Robb, S. A., and Roberts, R. G. (2001). Long mutant dystrophins and variable phenotypes: Evasion of nonsense-mediated decay? *Hum. Genet.* **109,** 402–407.

Kingston, H. M., Harper, P. S., Pearson, P. L., Davies, K. E., Williamson, R., and Page, D. (1983). Localisation of gene for Becker muscular dystrophy. *Lancet* **2,** 1200.

Knight, B., Laukaitis, C., Akhtar, N., Hotchin, N. A., Edlund, M., and Horwitz, A. R. (2000). Visualizing muscle cell migration *in situ*. *Curr. Biol.* **10,** 576–585.

Koenig, M., Hoffman, E. P., Bertelson, C. J., Monaco, A. P., Feener, C., and Kunkel, L. M. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50,** 509–517.

Koenig, M., Monaco, A. P., and Kunkel, L. M. (1988). The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* **53,** 219–226.

Koenig, M., Beggs, A. H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Muller, C. R., Lindlof, M., Kaariainen, H., de la Chapelle, A., Kiuru, A., Savontaus, M-L., Gilgenkrantz, H., Récan, D., Chelly, J., Kaplan, J-C., Covone, A. E., Archidiacono, N., Romeo, G., Liechti-Gallati, S., Schneider, V., Braga, S., Moser, H., Darras, B. T., Murphy, P., Francke, U., Chen, J. D., Morgan, G., Denton, M., Greenberg, C. R., Wrogemann, K., Blonden, L. A. J., van Paassen, H. M. B., van Ommen, G. J. B., and Kunkel, L. M. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: Correlation of severity with type of deletion. *Am. J. Hum. Genet.* **45,** 498–506.

Korenbaum, E., and Rivero, F. (2002). Calponin homology domains at a glance. *J. Cell Sci.* **115,** 3543–3545.

Kuhlman, P. A., Hemmings, L., and Critchley, D. R. (1992). The identification and characterisation of an actin-binding site in alpha-actinin by mutagenesis. *FEBS Lett.* **304,** 201–206.

Kurihara, H., Anderson, J. M., and Farquhar, M. G. (1995). Increased Tyr phosphorylation of ZO-1 during modification of tight junctions between glomerular foot processes. *Am. J. Physiol.* **268,** F514–524.

Law, R., Carl, P., Harper, S., Dalhaimer, P., Speicher, D. W., and Discher, D. E. (2003). Cooperativity in forced unfolding of tandem spectrin repeats. *Biophys. J.* **84,** 533–544.

Lazarides, E., and Burridge, K. (1975). Alpha-actinin: Immunoflourescent localisation of a muscle structural protein in nonmuscle cells. *Cell* **6,** 289–298.

Lebart, M. C., Mejean, C., Boyer, M., Roustan, C., and Benyamin, Y. (1990). Localization of a new alpha-actinin binding site in the COOH-terminal part of actin sequence. *Biochem. Biophys. Res. Commun.* **173,** 120–126.

Lebart, M. C., Mejean, C., Roustan, C., and Benyamin, Y. (1993). Further characterization of the alpha-actinin-actin interface and comparison with filamin-binding sites on actin. *J. Biol. Chem.* **268,** 5642–5648.

Lee, J. C., and Discher, D. E. (2001). Deformation-enhanced fluctuations in the red cell skeleton with theoretical relations to elasticity, connectivity, and spectrin unfolding. *Biophys. J.* **81,** 3178–3192.

Lemmon, M. A., and Ferguson, K. M. (1998). Pleckstrin homology domains. *Curr. Top. Microbiol. Immunol.* **228,** 39–74.

Lemmon, M. A., Ferguson, K. M., and Abrams, C. S. (2002). Pleckstrin homology domains and the cytoskeleton. *FEBS Lett.* **513,** 71–76.

Lenk, U., Oexle, K., Voit, T., Ancker, U., Hellner, K. A., Speer, A., and Hubner, C. (1996). A cysteine 3340 substitution in the dystroglycan-binding domain of dystrophin associated with Duchenne muscular dystrophy, mental retardation and absence of the ERG b-wave. *Hum. Mol. Genet.* **5,** 973–975.

Le Rumeur, E., Fichou, Y., Pottier, S., Gaboriau, F., Rondeau-Mouro, C., Vincent, M., Gallay, J., and Bondon, A. (2003). Interaction of dystrophin rod domain with

membrane phospholipids. Evidence of a close proximity between tryptophan residues and lipids. *J. Biol. Chem.* **278,** 5993–6001.

Levine, B. A., Moir, A. J., Patchell, V. B., and Perry, S. V. (1990). The interaction of actin with dystrophin. *FEBS Lett.* **263,** 159–162.

Levine, B. A., Moir, A. J., Patchell, V. B., and Perry, S. V. (1992). Binding sites involved in the interaction of actin with the N-terminal region of dystrophin. *FEBS Lett.* **298,** 44–48.

Li, X., and Bennett, V. (1996). Identification of the spectrin subunit and domains required for formation of spectrin/adducin/actin complexes. *J. Biol. Chem.* **271,** 15695–15702.

Liu, J., Taylor, D. W., and Taylor, K. A. (2004). A 3-D reconstruction of smooth muscle alpha-actinin by CryoEm reveals two different conformations at the actin-binding region. *J. Mol. Biol.* **338,** 115–125.

Loh, N. Y., Newey, S. E., Davies, K. E., and Blake, D. J. (2000). Assembly of multiple dystrobrevin-containing complexes in the kidney. *J. Cell Sci.* **113,** 2715–2724.

Love, D. R., Hill, D. F., Dickson, G., Spurr, N. K., Byth, B. C., Marsden, R. F., Walsh, F. S., Edwards, Y. H., and Davies, K. E. (1989). An autosomal transcript in skeletal muscle with homology to dystrophin. *Nature* **339,** 55–58.

Love, D. R., Flint, T. J., Genet, S. A., Middleton-Price, H. R., and Davies, K. E. (1991). Becker muscular dystrophy patient with a large intragenic dystrophin deletion: Implications for functional minigenes and gene therapy. *J. Med. Genet.* **28,** 860–864.

Lundberg, S., Lehto, V. P., and Backman, L. (1992). Characterization of calcium binding to spectrins. *Biochemistry* **31,** 5665–5671.

Lundberg, S., Bjork, J., Lofvenberg, L., and Backman, L. (1995). Cloning, expression and characterization of two putative calcium-binding sites in human non-erythroid alpha-spectrin. *Eur. J. Biochem.* **230,** 658–665.

Luther, P. K. (2000). Three-dimensional structure of a vertebrate muscle Z-band: Implications for titin and alpha-actinin binding. *J. Struct. Biol.* **129,** 1–16.

Lux, S. E., Tse, W. T., Menninger, J. C., John, K. M., Harris, P., Shalev, O., Chilcote, R. R., Marchesi, S. L., Watkins, P. C., Bennett, V., McIntosh, S., Collins, F. S., Francke, U., Ward, D. C., and Forget, B. G. (1990). Hereditary spherocytosis associated with deletion of human erythrocyte ankyrin gene on chromosome 8. *Nature* **345,** 736–739.

MacDonald, R. I., and Pozharski, E. V. (2001). Free energies of urea and of thermal unfolding show that two tandem repeats of spectrin are thermodynamically more stable than a single repeat. *Biochemistry* **40,** 3974–3984.

Maksymiw, R., Sui, S. F., Gaub, H., and Sackmann, E. (1987). Electrostatic coupling of spectrin dimers to phosphatidylserine containing lipid lamellae. *Biochemistry* **26,** 2983–2990.

Markle, D. R., Evans, E. A., and Hochmuth, R. M. (1983). Force relaxation and permanent deformation of erythrocyte membrane. *Biophys. J.* **42,** 91–98.

Masaki, T., Endo, M., and Ebashi, S. (1967). Localization of 6S component of a alpha-actinin at Z-band. *J. Biochem. (Tokyo)* **62,** 630–632.

Matsudaira, P. (1991). Modular organization of actin crosslinking proteins. *Trends Biochem. Sci.* **16,** 87–92.

Matsumura, K., Ervasti, J. M., Ohlendieck, K., Kahl, S. D., and Campbell, K. P. (1992). Association of dystrophin-related protein with dystrophin-associated proteins in mdx mouse muscle. *Nature* **360,** 588–591.

Matsumura, K., Yamada, H., Shimizu, T., and Campbell, K. P. (1993). Differential expression of dystrophin, utrophin and dystrophin-associated proteins in peripheral nerve. *FEBS Lett.* **334,** 281–285.

Mayer, B. J. (2001). SH3 domains: Complexity in moderation. *J. Cell Sci.* **114,** 1253–1263.

Mayer, B. J., Ren, R., Clark, K. L., and Baltimore, D. (1993). A putative modular domain present in diverse signaling proteins. *Cell* **73,** 629–630.

McGough, A. (1998). F-actin-binding proteins. *Curr. Opin. Struct. Biol.* **8,** 166–176.

McGough, A., Way, M., and DeRosier, D. (1994). Determination of the alpha-actinin-binding site on actin filaments by cryoelectron microscopy and image analysis. *J. Cell Biol.* **126,** 433–443.

Michalak, M., Fu, S. Y., Milner, R. E., Busaan, J. L., and Hance, J. E. (1996). Phosphorylation of the carboxyl-terminal region of dystrophin. *Biochem. Cell Biol.* **74,** 431–437.

Michaud, J. L., Lemieux, L. I., Dube, M., Vanderhyden, B. C., Robertson, S. J., and Kennedy, C. R. (2003). Focal and segmental glomerulosclerosis in mice with podocyte-specific expression of mutant alpha-actinin-4. *J. Am. Soc. Nephrol.* **14,** 1200–1211.

Mimura, N., and Asano, A. (1987). Further characterization of a conserved actin-binding 27-kDa fragment of actinogelin and alpha-actinins and mapping of their binding sites on the actin molecule by chemical cross-linking. *J. Biol. Chem.* **262,** 4717–4723.

Minetti, C., Tanji, K., and Bonilla, E. (1992). Immunologic study of vinculin in Duchenne muscular dystrophy. *Neurology* **42,** 1751–1754.

Monaco, A. P., Bertelson, C. J., Middlesworth, W., Colletti, C. A., Aldridge, J., Fischbeck, K. H., Bartlett, R., Pericak-Vance, M. A., Roses, A. D., and Kunkel, L. M. (1985). Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. *Nature* **316,** 842–845.

Moon, R. T., and Lazarides, E. (1983). Beta-spectrin limits alpha-spectrin assembly on membranes following synthesis in a chicken erythroid cell lysate. *Nature* **305,** 62–65.

Moores, C. A., and Kendrick-Jones, J. (2000a). Biochemical characterisation of the actin-binding properties of utrophin. *Cell Motil. Cytoskeleton* **46,** 116–128.

Moores, C. A., Keep, N. H., and Kendrick-Jones, J. (2000b). Structure of the utrophin actin-binding domain bound to F-actin reveals binding by an induced fit mechanism. *J. Mol. Biol.* **297,** 465–480.

Mukai, H., Toshimori, M., Shibata, H., Takanaga, H., Kitagawa, M., Miyahara, M., Shimakawa, M., and Ono, Y. (1997). Interaction of PKN with alpha-actinin. *J. Biol. Chem.* **272,** 4740–4746.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359,** 851–855.

Musacchio, A., Gibson, T., Rice, P., Thompson, J., and Saraste, M. (1993). The PH domain: A common piece in the structural patchwork of signalling proteins. *Trends Biochem. Sci.* **18,** 343–348.

Nakayama, S., and Kretsinger, R. H. (1994). Evolution of the EF-hand family of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **23,** 473–507.

Nguyen, T. M., Ellis, J. M., Love, D. R., Davies, K. E., Gatter, K. C., Dickson, G., and Morris, G. E. (1991). Localization of the DMDL gene-encoded dystrophin-related protein using a panel of nineteen monoclonal antibodies: Presence at neuromuscular junctions, in the sarcolemma of dystrophic skeletal muscle, in vascular and other smooth muscles, and in proliferating brain cell lines. *J. Cell Biol.* **115,** 1695–1700.

Noegel, A., Witke, W., and Schleicher, M. (1987). Calcium-sensitive non-muscle alpha-actinin contains EF-hand structures and highly conserved regions. *FEBS Lett.* **221,** 391–396.

Norwood, F. L., Sutherland-Smith, A. J., Keep, N. H., and Kendrick-Jones, J. (2000). The structure of the N-terminal actin-binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy. *Structure Fold Des.* **8,** 481–491.

O'Brien, K. F., and Kunkel, L. M. (2001). Dystrophin and muscular dystrophy: Past, present, and future. *Mol. Genet. Metab.* **74,** 75–88.

Otey, C. A., and Carpen, O. (2004). Alpha-actinin revisited: A fresh look at an old player. *Cell Motil. Cytoskeleton* **58,** 104–111.

Otey, C. A., Pavalko, F. M., and Burridge, K. (1990). An interaction between alpha-actinin and the beta 1 integrin subunit *in vitro. J. Cell Biol.* **111,** 721–729.

Otto, J. J. (1994). Actin-bundling proteins. *Curr. Opin. Cell Biol.* **6,** 105–109.

Palek, J., and Jarolim, P. (1993). Clinical expression and laboratory detection of red blood cell membrane protein mutations. *Semin. Hematol.* **30,** 249–283.

Pardo, J. V., Siliciano, J. D., and Craig, S. W. (1983). Vinculin is a component of an extensive network of myofibril-sarcolemma attachment regions in cardiac muscle fibers. *J. Cell Biol.* **97,** 1081–1088.

Pascual, J., Castresana, J., and Saraste, M. (1997). Evolution of the spectrin repeat. *Bioessays* **19,** 811–817.

Pasternak, C., Wong, S., and Elson, E. L. (1995). Mechanical function of dystrophin in muscle cells. *J. Cell Biol.* **128,** 355–361.

Pearce, M., Blake, D. J., Tinsley, J. M., Byth, B. C., Campbell, L., Monaco, A. P., and Davies, K. E. (1993). The utrophin and dystrophin genes share similarities in genomic structure. *Hum. Mol. Genet.* **2,** 1765–1772.

Pekrun, A., Eber, S. W., Kuhlmey, A., and Schroter, W. (1993). Combined ankyrin and spectrin deficiency in hereditary spherocytosis. *Ann. Hematol.* **67,** 89–93.

Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M., and Sweeney, H. L. (1993). Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proc. Natl. Acad. Sci. USA* **90,** 3710–3714.

Phelps, S. F., Hauser, M. A., Cole, N. M., Rafael, J. A., Hinkle, R. T., Faulkner, J. A., and Chamberlain, J. S. (1995). Expression of full-length and truncated dystrophin minigenes in transgenic mdx mice. *Hum. Mol. Genet.* **4,** 1251–1258.

Pollard, T. D. (1984). Purification of a high molecular weight actin filament gelation protein from Acanthamoeba that shares antigenic determinants with vertebrate spectrins. *J. Cell Biol.* **99,** 1970–1980.

Pons, F., Robert, A., Fabbrizio, E., Hugon, G., Califano, J. C., Fehrentz, J. A., Martinez, J., and Mornet, D. (1994). Utrophin localization in normal and dystrophin-deficient heart. *Circulation* **90,** 369–374.

Ponting, C. P., Blake, D. J., Davies, K. E., Kendrick-Jones, J., and Winder, S. J. (1996). ZZ and TAZ: New putative zinc fingers in dystrophin and other proteins. *Trends Biochem. Sci.* **21,** 11–13.

Porter, G. A., Dmytrenko, G. M., Winkelmann, J. C., and Bloch, R. J. (1992). Dystrophin colocalizes with beta-spectrin in distinct subsarcolemmal domains in mammalian skeletal muscle. *J. Cell Biol.* **117,** 997–1005.

Prior, T. W., Bartolo, C., Pearl, D. K., Papp, A. C., Snyder, P. J., Sedra, M. S., Burghes, A. H., and Mendell, J. R. (1995). Spectrum of small mutations in the dystrophin coding region. *Am. J. Hum. Genet.* **57,** 22–33.

Raae, A. J., Banuelos, S., Ylanne, J., Olausson, T., Goldie, K. N., Wendt, T., Hoenger, A., and Saraste, M. (2003). Actin binding of a minispectrin. *Biochim. Biophys. Acta* **1646,** 67–76.

Raats, C. J., van den Born, J., Bakker, M. A., Oppers-Walgreen, B., Pisa, B. J., Dijkman, H. B., Assmann, K. J., and Berden, J. H. (2000). Expression of agrin, dystroglycan, and utrophin in normal renal tissue and in experimental glomerulopathies. *Am. J. Pathol.* **156,** 1749–1765.

Rafael, J. A., Cox, G. A., Corrado, K., Jung, D., Campbell, K. P., and Chamberlain, J. S. (1996). Forced expression of dystrophin deletion constructs reveals structure-function correlations. *J. Cell Biol.* **134,** 93–102.

Rando, T. A. (2001). The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. *Muscle Nerve* **24,** 1575–1594.

Rebecchi, M. J., and Scarlata, S. (1998). Pleckstrin homology domains: A common fold with diverse functions. *Annu. Rev. Biophys. Biomol. Struct.* **27,** 503–528.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259,** 1157–1161.

Rentschler, S., Linn, H., Deininger, K., Bedford, M. T., Espanel, X., and Sudol, M. (1999). The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan. *Biol. Chem.* **380,** 431–442.

Rief, M., Pascual, J., Saraste, M., and Gaub, H. E. (1999). Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *J. Mol. Biol.* **286,** 553–561.

Roberts, R. G., Coffey, A. J., Bobrow, M., and Bentley, D. R. (1993). Exon structure of the human dystrophin gene. *Genomics* **16,** 536–538.

Roberts, R. G., Gardner, R. J., and Bobrow, M. (1994). Searching for the 1 in 2,400,000: A review of dystrophin gene point mutations. *Hum. Mutat.* **4,** 1–11.

Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H., and Lewis, S. (2000). Comparative genomics of the eukaryotes. *Science* **287,** 2204–2215.

Rybakova, I. N., Patel, J. R., and Ervasti, J. M. (2000). The dystrophin complex forms a mechanically strong link between the sarcolemma and costameric actin. *J. Cell Biol.* **150,** 1209–1214.

Rybakova, I. N., Patel, J. R., Davies, K. E., Yurchenco, P. D., and Ervasti, J. M. (2002). Utrophin binds laterally along actin filaments and can couple costameric actin with sarcolemma when overexpressed in dystrophin-deficient muscle. *Mol. Biol. Cell* **13,** 1512–1521.

Rybicki, A. C., Heath, R., Wolf, J. L., Lubin, B., and Schwartz, R. S. (1988). Deficiency of protein 4.2 in erythrocytes from a patient with a Coombs negative hemolytic anemia. Evidence for a role of protein 4.2 in stabilizing ankyrin on the membrane. *J. Clin. Invest.* **81,** 893–901.

Sahr, K. E., Laurila, P., Kotula, L., Scarpa, A. L., Coupal, E., Leto, T. L., Linnenbach, A. J., Winkelmann, J. C., Speicher, D. W., Marchesi, V. T., Curtis, P. J., and Forget,

B. G. (1990). The complete cDNA and polypeptide sequences of human erythroid alpha-spectrin. *J. Biol. Chem.* **265,** 4434–4443.

Salmikangas, P., Mykkanen, O. M., Gronholm, M., Heiska, L., Kere, J., and Carpen, O. (1999). Myotilin, a novel sarcomeric protein with two Ig-like domains, is encoded by a candidate gene for limb-girdle muscular dystrophy. *Hum. Mol. Genet.* **8,** 1329–1336.

Savvides, P., Shalev, O., John, K. M., and Lux, S. E. (1993). Combined spectrin and ankyrin deficiency is common in autosomal dominant hereditary spherocytosis. *Blood* **82,** 2953–2960.

Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997). Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly. *Adv. Protein Chem.* **50,** 61–122.

Shirato, I., Hosser, H., Kimura, K., Sakai, T., Tomino, Y., and Kriz, W. (1996). The development of focal segmental glomerulosclerosis in masugi nephritis is based on progressive podocyte damage. *Virchows Arch.* **429,** 255–273.

Smoyer, W. E., Mundel, P., Gupta, A., and Welsh, M. J. (1997). Podocyte alpha-actinin induction precedes foot process effacement in experimental nephrotic syndrome. *Am. J. Physiol.* **273,** F150–157.

Speicher, D. W., and Marchesi, V. T. (1984). Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature* **311,** 177–180.

Stradal, T., Kranewitter, W., Winder, S. J., and Gimona, M. (1998). CH domains revisited. *FEBS Lett.* **431,** 134–137.

Straub, V., and Campbell, K. P. (1997). Muscular dystrophies and the dystrophin-glycoprotein complex. *Curr. Opin. Neurol.* **10,** 168–175.

Straub, V., Bittner, R. E., Leger, J. J., and Voit, T. (1992). Direct visualization of the dystrophin network on skeletal muscle fiber membrane. *J. Cell Biol.* **119,** 1183–1191.

Sutherland-Smith, A. J., Moores, C. A., Norwood, F. L., Hatch, V., Craig, R., Kendrick-Jones, J., and Lehman, W. (2003). An atomic model for actin binding by the CH domains and spectrin-repeat modules of utrophin and dystrophin. *J. Mol. Biol.* **329,** 15–33.

Tang, J., Taylor, D. W., and Taylor, K. A. (2001). The three-dimensional structure of alpha-actinin obtained by cryoelectron microscopy suggests a model for Ca(2+)-dependent actin binding. *J. Mol. Biol.* **310,** 845–858.

Taylor, K. A., and Taylor, D. W. (1993). Projection image of smooth muscle alpha-actinin from two-dimensional crystals formed on positively charged lipid layers. *J. Mol. Biol.* **230,** 196–205.

Taylor, K. A., Taylor, D. W., and Schachat, F. (2000). Isoforms of alpha-actinin from cardiac, smooth, and skeletal muscle form polar arrays of actin filaments. *J. Cell Biol.* **149,** 635–646.

Thomas, G. H., Newbern, E. C., Korte, C. C., Bales, M. A., Muse, S. V., Clark, A. G., and Kiehart, D. P. (1997). Intragenic duplication and divergence in the spectrin superfamily of proteins. *Mol. Biol. Evol.* **14,** 1285–1295.

Tinsley, J. M., Blake, D. J., Roche, A., Fairbrother, U., Riss, J., Byth, B. C., Knight, A. E., Kendrick-Jones, J., Suthers, G. K., Love, D. R., Edwards, Y. H., and Davies, K. E. (1992). Primary structure of dystrophin-related protein. *Nature* **360,** 591–593.

Tommasi di Vignano, A., Di Zenzo, G., Sudol, M., Cesareni, G., and Dente, L. (2000). Contribution of the different modules in the utrophin carboxy-terminal region to the formation and regulation of the DAP complex. *FEBS Lett.* **471,** 229–234.

Trave, G., Lacombe, P. J., Pfuhl, M., Saraste, M., and Pastore, A. (1995). Molecular mechanism of the calcium-induced conformational change in the spectrin EF-hands. *EMBO J.* **14,** 4922–4931.

Tufty, R. M., and Kretsinger, R. H. (1975). Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. *Science* **187,** 167–169.

Vainzof, M., Takata, R. I., Passos-Bueno, M. R., Pavanello, R. C., and Zatz, M. (1993). Is the maintenance of the C-terminus domain of dystrophin enough to ensure a milder Becker muscular dystrophy phenotype? *Hum. Mol. Genet.* **2,** 39–42.

Viel, A. (1999). Alpha-actinin and spectrin structures: An unfolding family story. *FEBS Lett.* **460,** 391–394.

Viel, A., and Branton, D. (1996). Spectrin: On the path from structure to function. *Curr. Opin. Cell Biol.* **8,** 49–55.

Virel, A., and Backman, L. (2004). Molecular evolution and structure of alpha-actinin. *Mol. Biol. Evol.* **21,** 1024–1031.

Wang, D. S., and Shaw, G. (1995). The association of the C-terminal region of beta I sigma II spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1,4,5 triphosphate binding site. *Biochem. Biophys. Res. Commun.* **217,** 608–615.

Wang, D. S., Miller, R., Shaw, R., and Shaw, G. (1996). The pleckstrin homology domain of human beta I sigma II spectrin is targeted to the plasma membrane *in vivo. Biochem. Biophys. Res. Commun.* **225,** 420–426.

Wasenius, V. M., Saraste, M., Salven, P., Eramaa, M., Holm, L., and Lehto, V. P. (1989). Primary structure of the brain alpha-spectrin. *J. Cell Biol.* **108,** 79–93.

Waugh, R., and Evans, E. A. (1979). Thermoelasticity of red blood cell membrane. *Biophys. J.* **26,** 115–131.

Waugh, R. E., and Agre, P. (1988). Reductions of erythrocyte membrane viscoelastic coefficients reflect spectrin deficiencies in hereditary spherocytosis. *J. Clin. Invest.* **81,** 133–141.

Way, M., Pope, B., and Weeds, A. G. (1992). Evidence for functional homology in the F-actin binding domains of gelsolin and alpha-actinin: Implications for the requirements of severing and capping. *J. Cell Biol.* **119,** 835–842.

Wells, D. J., Wells, K. E., Asante, E. A., Turner, G., Sunada, Y., Campbell, K. P., Walsh, F. S., and Dickson, G. (1995). Expression of human full-length and minidystrophin in transgenic mdx mice: Implications for gene therapy of Duchenne muscular dystrophy. *Hum. Mol. Genet.* **4,** 1245–1250.

Wessel, G. M., and Chen, S. W. (1993). Transient, localized accumulation of alpha-spectrin during sea urchin morphogenesis. *Dev. Biol.* **155,** 161–171.

Williams, M. W., and Bloch, R. J. (1999). Extensive but coordinated reorganization of the membrane skeleton in myofibers of dystrophic (mdx) mice. *J. Cell Biol.* **144,** 1259–1270.

Winder, S. J. (1996). Structure-function relationships in dystrophin and utrophin. *Biochem. Soc. Trans.* **24,** 497–501.

Winder, S. J. (1997). The membrane-cytoskeleton interface:The role of dystrophin and utrophin. *J. Muscle Res. Cell Motil.* **18,** 617–629.

Winder, S. J. (2001). The complexities of dystroglycan. *Trends Biochem. Sci.* **26,** 118–124.

Winder, S. J., and Walsh, M. P. (1990). Structural and functional characterization of calponin fragments. *Biochem. Int.* **22,** 335–341.

Winder, S. J., Gibson, T. J., and Kendrick-Jones, J. (1995a). Dystrophin and utrophin: The missing links!. *FEBS Lett.* **369,** 27–33.

Winder, S. J., Hemmings, L., Maciver, S. K., Bolton, S. J., Tinsley, J. M., Davies, K. E., Critchley, D. R., and Kendrick-Jones, J. (1995b). Utrophin actin binding domain: Analysis of actin binding and cellular targeting. *J. Cell Sci.* **108,** 63–71.

Winkelmann, J. C., and Forget, B. G. (1993). Erythroid and nonerythroid spectrins. *Blood* **81,** 3173–3185.

Winnard, A. V., Klein, C. J., Coovert, D. D., Prior, T., Papp, A., Snyder, P., Bulman, D. E., Ray, P. N., McAndrew, P., and King, W. (1993). Characterization of translational frame exception patients in Duchenne/Becker muscular dystrophy. *Hum. Mol. Genet.* **2,** 737–744.

Wolfe, L. C., John, K. M., Falcone, J. C., Byrne, A. M., and Lux, S. E. (1982). A genetic defect in the binding of protein 4.1 to spectrin in a kindred with hereditary spherocytosis. *N. Engl. J. Med.* **307,** 1367–1374.

Woods, C. M., and Lazarides, E. (1985). Degradation of unassembled alpha- and beta-spectrin by distinct intracellular pathways: Regulation of spectrin topogenesis by beta-spectrin degradation. *Cell* **40,** 959–969.

Xia, H., Winokur, S. T., Kuo, W. L., Altherr, M. R., and Bredt, D. S. (1997). Actinin-associated LIM protein: Identification of a domain interaction between PDZ and spectrin-like repeat motifs. *J. Cell Biol.* **139,** 507–515.

Yao, J., Le, T. C., Kos, C. H., Henderson, J. M., Allen, P. G., Denker, B. M., and Pollak, M. R. (2004). Alpha-actinin-4-mediated FSGS: An inherited kidney disease caused by an aggregated and rapidly degraded cytoskeletal protein. *PLoS Biol.* **2,** 787–794.

Ylanne, J., Scheffzek, K., Young, P., and Saraste, M. (2001a). Crystal structure of the alpha-actinin rod reveals an extensive torsional twist. *Structure* **9,** 597–604.

Ylanne, J., Scheffzek, K., Young, P., and Saraste, M. (2001b). Crystal structure of the alpha-actinin rod: Four spectrin repeats forming a tight dimer. *Cell Mol. Biol. Lett.* **6,** 234.

Young, P., Ferguson, C., Banuelos, S., and Gautel, M. (1998). Molecular structure of the sarcomeric Z-disk: Two types of titin interactions lead to an asymmetrical sorting of alpha-actinin. *EMBO J.* **17,** 1614–1624.

Zarrinpar, A., and Lim, W. A. (2000). Converging on proline: The mechanism of WW domain peptide recognition. *Nat. Struct. Biol.* **7,** 611–613.

Zhang, M., Tanaka, T., and Ikura, M. (1995a). Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.* **2,** 758–767.

Zhang, P., Talluri, S., Deng, H., Branton, D., and Wagner, G. (1995b). Solution structure of the pleckstrin homology domain of *Drosophila* beta-spectrin. *Structure* **3,** 1185–1195.

Zhou, M. M., Ravichandran, K. S., Olejniczak, E. F., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature* **378,** 584–592.

FIBRINOGEN AND FIBRIN

By JOHN W. WEISEL

Department of Cell and Developmental Biology, University of Pennsylvania
School of Medicine, Philadelphia, Pennsylvania 19104-6058

I. Introduction 248
II. Structure and Properties of Fibrinogen 251
A. Physico-Chemical Properties, Amino Acid Sequence, and Disulfide Bonding 251
B. Domains 253
C. Coiled Coils 256
D. Carbohydrate 256
E. X-ray Crystallographic Structure 257
F. Calcium Binding 259
III. Biosynthesis and Metabolism of Fibrinogen 260
IV. Conversion of Fibrinogen to Fibrin 263
A. Thrombin-Catalyzed Cleavage of Fibrinopeptides 263
B. Polymerization Steps 264
C. Binding Pockets 269
D. Other Binding Interactions 269
E. Covalent Crosslinking of Fibrin by Factor XIIIa 270
F. Mechanical Properties of Clots 272
V. Binding of Other Proteins to Fibrin(ogen) 273
A. Fibronectin and Albumin 273
B. Thrombospondin, von Willebrand Factor, and Fibulin 274
C. Fibroblast Growth Factor-2, Vascular Endothelial Growth Factor, and Interleukin-1 274
VI. Fibrinolysis 275
VII. Fibrinogen Binding to Integrins in Platelet Aggregation and Other Cellular Interactions 277
VIII. Dysfibrinogenemias, Variants, and Heterogeneity of Fibrinogen 279
IX. Evolution of Fibrinogen and Fibrinogen-Like Domains 283
X. Conclusions 284
References 285

Abstract

Fibrinogen is a large, complex, fibrous glycoprotein with three pairs of polypeptide chains linked together by 29 disulfide bonds. It is 45 nm in length, with globular domains at each end and in the middle connected by α-helical coiled-coil rods. Both strongly and weakly bound calcium ions are important for maintenance of fibrinogen's structure and functions. The fibrinopeptides, which are in the central region, are cleaved by thrombin to convert soluble fibrinogen to insoluble fibrin polymer, via

DOI: 10.1016/S0065-3233(04)70008-X
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

intermolecular interactions of the ''knobs'' exposed by fibrinopeptide removal with ''holes'' always exposed at the ends of the molecules. Fibrin monomers polymerize via these specific and tightly controlled binding interactions to make half-staggered oligomers that lengthen into protofibrils. The protofibrils aggregate laterally to make fibers, which then branch to yield a three-dimensional network—the fibrin clot—essential for hemostasis. X-ray crystallographic structures of portions of fibrinogen have provided some details on how these interactions occur. Finally, the transglutaminase, Factor XIIIa, covalently binds specific glutamine residues in one fibrin molecule to lysine residues in another via isopeptide bonds, stabilizing the clot against mechanical, chemical, and proteolytic insults. The gene regulation of fibrinogen synthesis and its assembly into multichain complexes proceed via a series of well-defined steps. Alternate splicing of two of the chains yields common variant molecular isoforms. The mechanical properties of clots, which can be quite variable, are essential to fibrin's functions in hemostasis and wound healing. The fibrinolytic system, with the zymogen plasminogen binding to fibrin together with tissue-type plasminogen activator to promote activation to the active enzyme plasmin, results in digestion of fibrin at specific lysine residues. Fibrin(ogen) also specifically binds a variety of other proteins, including fibronectin, albumin, thrombospondin, von Willebrand factor, fibulin, fibroblast growth factor-2, vascular endothelial growth factor, and interleukin-1. Studies of naturally occurring dysfibrinogenemias and variant molecules have increased our understanding of fibrinogen's functions. Fibrinogen binds to activated αIIbβ3 integrin on the platelet surface, forming bridges responsible for platelet aggregation in hemostasis, and also has important adhesive and inflammatory functions through specific interactions with other cells. Fibrinogen-like domains originated early in evolution, and it is likely that their specific and tightly controlled intermolecular interactions are involved in other aspects of cellular function and developmental biology.

I. Introduction

Fibrinogen is a fibrous protein that was first classified with keratin, myosin, and epidermin based on its 5.1 Å repeat in wide-angle X-ray diffraction patterns (Bailey *et al.*, 1943), which was later discovered to be associated with the α-helical coiled-coil structure. It is a glycoprotein normally present in human blood plasma at a concentration of about 2.5 g/L and is essential for hemostasis, wound healing, inflammation, angiogenesis, and other biological functions. It is a soluble macromolecule, but forms a clot or insoluble gel on conversion to fibrin by the action of the

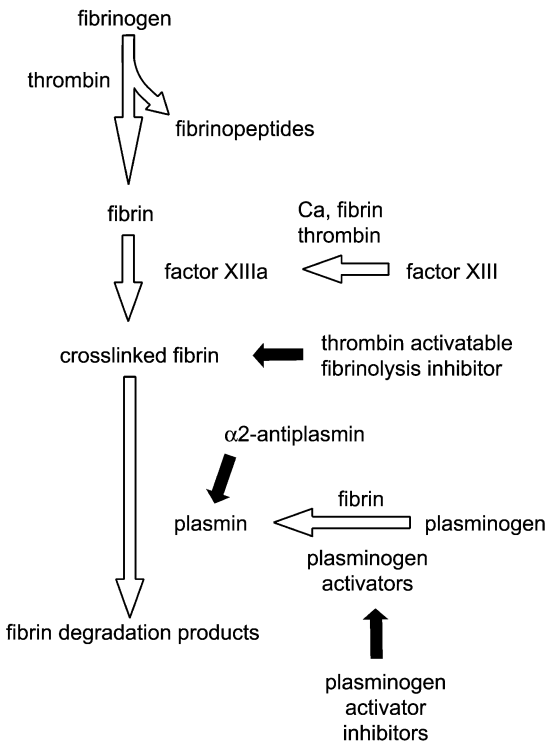


FIG. 1. Basic scheme of fibrin polymerization and fibrinolysis. The clot is formed on the conversion of fibrinogen to fibrin by cleavage of the fibrinopeptides by thrombin, followed by stabilization of the network with isopeptide bonds by the transglutaminase Factor XIIIa. The clot is dissolved through proteolysis by the enzyme plasmin, which is activated on the fibrin surface by plasminogen activators. This process is controlled by several inhibitory reactions (black arrows).

serine proteolytic enzyme thrombin (Fig. 1), which itself is activated by a cascade of enzymatic reactions triggered by injury or a foreign surface. A mechanically stable, protease-resistant clot is necessary to prevent blood loss and promote wound healing.

Fibrinogen is also necessary for the aggregation of blood platelets, an initial step in hemostasis. Each end of a fibrinogen molecule can bind with high affinity to the integrin receptor on activated platelets, αIIbβ3, so that the bifunctional fibrinogen molecules act as bridges to link platelets. In its various functions as a clotting and adhesive protein, the fibrinogen

molecule is involved in many intermolecular interactions and has specific binding sites for several proteins and cells.

Fibrin clots are dissolved by another series of enzymatic reactions termed the *fibrinolytic system* (Fig. 1). The proenzyme plasminogen is activated to plasmin by a specific proteolytic enzyme, typically tissue-type plasminogen activator or urokinase-type plasminogen activator. Plasmin then cleaves fibrin at certain unique locations to dissolve the clot. The activation of the fibrinolytic system is greatly enhanced by taking place on the surface of fibrin. Thus, these reactions are highly specific for cleavage of the insoluble fibrin clot, rather than circulating fibrinogen.

There is a dynamic equilibrium between clotting and fibrinolysis, so that the conversion of fibrinogen to fibrin and the dissolution of the clot must be carefully regulated. Any imbalance can result in either loss of blood from hemorrhage or blockage of the flow of blood through a vessel from thrombosis. Thrombosis, often accompanying atherosclerosis or other pathological processes, is the most common cause of myocardial infarction and stroke. As a consequence, various fibrinolytics that activate plasminogen are now commonly used clinically to treat these conditions.

The clinical significance of fibrin was first recognized by Virchow. In 1859, Deni de Commercy proposed the existence of a precursor of fibrin that he called fibrinogen (Blombäck, 2001). Although Hammarsten first purified fibrinogen from horse plasma in 1876 (Blombäck, 2001), human fibrinogen was isolated in large quantities only fifty years ago (Cohn *et al.*, 1946) and has been extensively studied since then. Fibrinogen and the other components of the clotting system are commonly isolated from human blood plasma, using purification methods based on fibrinogen's low solubility in various solvents or its isoelectric point (Blombäck and Blombäck, 1956; Cohn *et al.*, 1946; Laki, 1951).

Fibrinogen was last reviewed in this series in 1973 (Doolittle, 1973). However, many other reviews have appeared in the meantime, which have summarized various aspects of the biology and biochemistry of fibrinogen and fibrin (Blombäck, 1991; Budzynski, 1986; Doolittle, 1984; Greenberg *et al.*, 2003; Hantgan *et al.*, 2000; Henschen and McDonagh, 1986; Mosesson *et al.*, 2001; Shafer and Higgins, 1988; Tooney and Weisel, 1979; Weisel and Cederholm-Williams, 1997). There has been an impressive amount of research on all aspects of fibrinogen and fibrin in the past 30 years. This Chapter will provide an outline of some of this body of work with selected references to the relevant literature.

As one indication of the progress, there have been generally accepted resolutions of the three fundamental controversies cited by Doolittle (1973), although of course not everyone agrees with all of these conclusions: (1) while there was then ''no general accord on the shape of the

TABLE I
Physicochemical Characteristics of Human Fibrinogen

Molecular mass	340,000
Molecular volume	3.7×10^3 nm^3
Sedimentation coefficient ($S_{20,w}$)	7.8×10^{-13} s
Translational diffusion coefficient ($D_{20,w}$)	1.9×10^{-7}
Rotary diffusion coefficient ($O_{20,w}$)	40,000 s^{-1}
Frictional ratio (f/f_0)	2.34
Partial specific volume	0.72 cm^3/g
Extinction coefficient ($A_{280,\,1\%}$)	15.1
Intrinsic viscosity (η)	0.25 dl/g
Degree of hydration (g/g of protein)	6
α-helix content	33%
Isoelectric point	5.5

fibrinogen molecule,'' first electron microscopy and then X-ray crystallography have been used for the determination of the detailed structure of fibrinogen; (2) while ''the nature of the forces holding subunit portions of fibrinogen together'' was not known, biochemical and structural studies showed the bonds involved; and (3) while ''the general location of the fibrinopeptides in the parent molecule'' was undecided at that time, now we know where they are in the molecule, although they have not been resolved at atomic resolution. Moreover, we now know a great deal about the process of fibrin polymerization, fibrinogen synthesis, and interactions with plasma proteins and cells.

II. STRUCTURE AND PROPERTIES OF FIBRINOGEN

A. *Physico-Chemical Properties, Amino Acid Sequence, and Disulfide Bonding*

The physicochemical characteristics of fibrinogen reflect its nature as a fibrous protein (Table I). Fibrinogen is a large glycoprotein made up of three pairs of polypeptide chains, designated Aα, Bβ, and γ, with molecular masses of 66,500, 52,000, and 46,500 Da, respectively (Fig. 2). The posttranslational addition of asparagine-linked carbohydrate to the Bβ and γ chains brings the total molecular mass to about 340,000 Da. The nomenclature for fibrinogen, $(A\alpha, B\beta, \gamma)_2$, arises from the designation of the small peptides that are cleaved from fibrinogen by thrombin to yield fibrin as fibrinopeptides A and B and the parent chains, without the

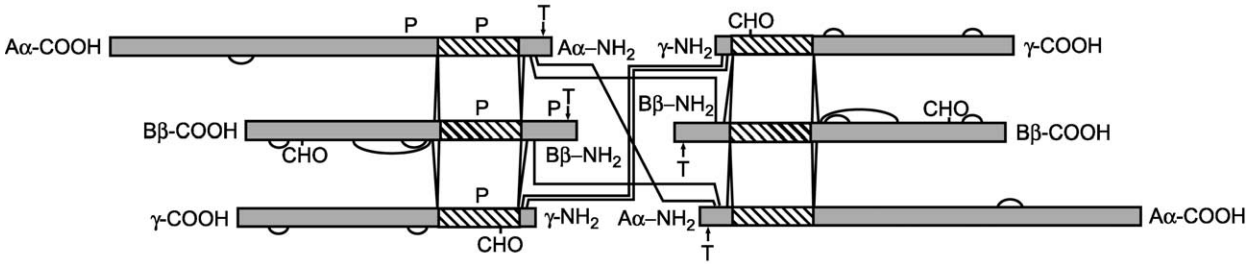


FIG. 2. Schematic diagram of the three pairs of polypeptide chains of fibrinogen. The Aα, Bβ, and γ chains are represented by bars with lengths proportional to the number of amino acid residues in each chain and the N- and C-terminal ends of the chains are labeled. The coiled-coil regions are indicated by the diagonally striped boxes, while the intra- and interchain disulfide bonds are indicated by solid lines. Carbohydrate attachment sites are labeled with CHO, while thrombin and major plasmin cleavage sites are indicated by T and P, respectively. (Adapted from Fig. 12-1 of Hantgan *et al.*, 2000.)

fibrinopeptides, as α and β. No peptides are cleaved from the γ chains by thrombin.

The entire amino acid sequence of all polypeptide chains of human fibrinogen has been determined by the methods of protein chemistry and later deduced from the nucleotide sequences of the cDNA coding for the polypeptide chains (Doolittle, 1984; Henschen and McDonagh, 1986). There are 610, 461, and 411 amino acids in each of the common forms of the human Aα, Bβ, and γ chains, respectively. The amino acid sequences of the three chains are homologous, but important differences exist, giving rise to specific functions for certain molecular domains.

All six chains are held together by 29 disulfide bonds (Henschen and McDonagh, 1986) to make two half molecules. There are 8, 11, and 10 cysteine residues in the Aα, Bβ, and γ chains, respectively. The amino termini of all six chains are held together by disulfide bonds in the central domain (Fig. 2). Three interchain disulfide bonds link the two halves of the molecule together, one between the two Aα chains and two between the two γ chains. The A and B fibrinopeptides are also located at the N-terminal ends of the α and β chains, respectively. A single interchain disulfide bond connects the Aα and Bβ chains within each half molecule. Unusual Cys-Pro-X-X-Cys sequences occurring twice in each chain are involved in disulfide ring structures, in which all three chains are linked together at each end of α-helical coiled coils (Doolittle, 1984). The remainder of the Aα chain contains one intrachain disulfide, while the Bβ chain contains three intrachain disulfides and the γ chain contains two.

B. Domains

The shape of fibrinogen and its organization into domains, or independently folded units, have been defined by a variety of physicochemical and structural techniques. Fibrinogen was one of the first biological macromolecules to be visualized by electron microscopy (Hall and Slayter, 1959). Fibrinogen rotary shadowed with platinum was seen to be an elongated molecule about 46 nm in length with nodular regions at each end and in the middle connected by rodlike strands. Subsequent studies confirmed this model, although many conflicting models were proposed in the intervening years. Subdivision of the end nodules was demonstrated and the αC domains were localized (Fig. 3A; Erickson and Fowler, 1983; Mosesson *et al.*, 1981; Weisel *et al.*, 1981, 1985; Williams, 1983).

The central region contains the two pairs of A and B fibrinopeptides and binds thrombin. The distal end nodule is the C-terminal γ chain (γC

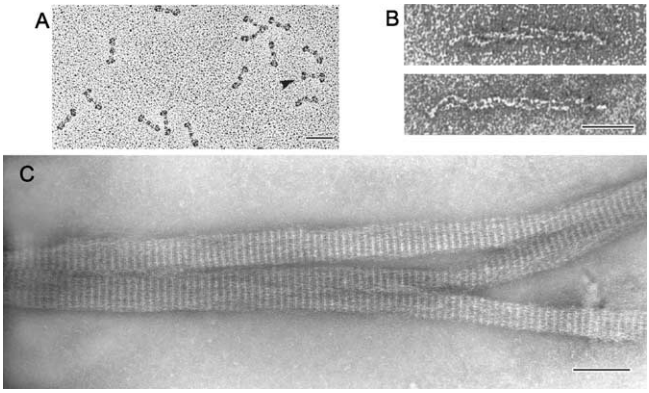


FIG. 3. Transmission electron microscope images of fibrinogen molecules, protofibrils, and fibers. (A) Fibrinogen molecules rotary shadowed with a thin layer of metal, showing the 45–47 nm long elongated protein molecules, with a central domain and two nodules at each end, connected by thin rods. Magnification bar, 50 nm. (Image from Weisel *et al.*, 1985). (B) Negatively contrasted fibrin protofibrils, showing the half staggering of monomers and the twisting of two filaments of the protofibril. In negative contrast, protein excludes stain and these areas appear bright. Magnification bar, 50 nm. (Image from Medved *et al.*, 1990) (C) Negatively contrasted fibrin fibers. Fibers have a repeat of 22.5 nm, with a distinctive band pattern that arises because areas with high protein density exclude stain and appear bright, while areas with lower protein density allow more stain to penetrate and appear darker. Note the trifunctional branch points, at which a portion of the fiber diverges from the remainder. The striations of the fibrin band pattern are aligned at the branch point. Magnification bar, 200 nm. (Image from Weisel, 2004b.)

nodule) and is made up of three domains, while the proximal end nodule is the C-terminal Bβ chain (βC nodule), also made up of three domains (Medved, 1990; Rao *et al.*, 1991; Weisel *et al.*, 1985). The C-terminal portions of the two Aα chains (αC domains) extend from the molecular ends and interact with each other and with the central domains in fibrinogen (Gorkun *et al.*, 1994; Veklich *et al.*, 1993).

Differential scanning microcalorimetry has been used together with limited proteolysis to identify cooperative domains in fibrinogen. It was found that each terminal part of the fibrinogen molecule is made up of six cooperative domains, with three domains formed by the C-terminal part of the β-chain and three other domains formed by the C-terminal part of the γ-chain (Medved, 1990; Medved *et al.*, 1986, 1997; Privalov and Medved, 1982). The amino acid sequence of the αC domains suggested that each is made up of a globular and an extended portion. Microcalorimetry

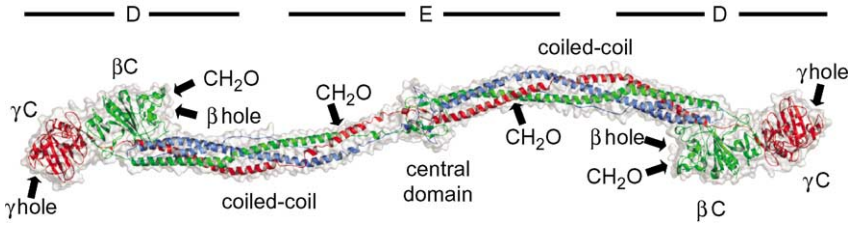


FIG. 4. X-ray crystallographic structure of fibrinogen. The Aα chains are blue, while the Bβ chains are green, and the γ chains are red. Most of the C-terminal portion of the α chains is missing. The extent of fragments D and E, created by plasmin cleavage in the middle of the coiled coil, are indicated. The central domain is connected to the end regions via α-helical coiled-coil rodlike regions. The βC and γC nodules contain the holes that are complementary to the knobs in the central nodule, although the knobs uncovered by cleavage of the fibrinopeptides are not seen because they are disordered and/or flexible. Carbohydrate attachment sites are indicated by CH_2O. The atomic resolution structure of fibrinogen has been built up from X-ray crystallographic studies of human γC module and fragment D, bovine fragment E, modified bovine fibrinogen, and chicken fibrinogen. Where comparisons are available, it seems that there are few differences in structures of the fibrinogens from different species (Brown *et al.*, 2000; Everse *et al.*, 1998; Madrazo *et al.*, 2001; Yang *et al.*, 2001). (This model was kindly generated by Igor Pechik using atomic coordinates PDB ID 1M1J, which replaced the original entry 1JFE (Yang *et al.*, 2001).)

confirmed this result and showed that the two αC domains interact intramolecularly.

Several proteolytic fragments of fibrinogen have been important for defining its domain structure and identifying functional sites (Hantgan *et al.*, 2000). The fibrinolytic enzyme plasmin dissolves the fibrin clot, but also cleaves fibrinogen at similar locations. Initial digestion removes the C-terminal Aα chains and Bβ 1-42, creating fragment X. Cleavages in all three chains then yield two D fragments and one E fragment (Fig. 4). Incomplete cleavage can produce a single D fragment plus a Y fragment (D + E). Through a variety of studies, the organization of each of these fragments has been defined. Fragment E contains the central region of fibrinogen and about half of each coiled coil extending on either side. Each fragment D consists of the remainder of one coiled-coil rod and the C-terminal portion of the Bβ and γ chains plus part of the Aα chain. The prevalence of studies using these fragments has led to extensive use of descriptive terms such as *D domain* or *E domain* or similar nomenclature, but it is important to note that these fragments are not domains; instead, each is made up of multiple domains, so they should be called *D regions* or *E regions*.

C. Coiled Coils

The central and end regions of fibrinogen are joined together by α-helical coiled-coil rods (see Chapter 3). There are 111 or 112 amino acid residues from each of the Aα, Bβ, and γ chains that are bordered by disulfide rings, as described above. Amino acid sequence predictions suggested the presence of coiled-coil structures that would account for the α structure first identified in wide-angle X-ray fiber diagrams of fibrinogen (Doolittle *et al.*, 1978; Parry, 1978). In the coiled coil, most nonpolar residues point inward to form a hydrophobic core, while charged residues are exposed to solvent on the outside. Furthermore, these predictions suggested that this coiled-coil structure was interrupted in the middle in a region known to be susceptible to plasmin cleavage, with some other possible interruptions. The X-ray crystallographic structures of fragments D and E and of the entire fibrinogen molecule (described below) showed the coiled coil and revealed that there is a fourth strand for part of the coiled coil, formed from the Aα chain that is folded back after the distal disulfide ring.

D. Carbohydrate

Four biantennary oligosaccharide chains are linked to each molecule of fibrinogen by way of N-glycosyl bonds at Bβ364 and γ52. These carbohydrate attachment sites contain the NXS or NXT sequences that are typical of N-glycosylation. On the other hand, the Aα chains are devoid of any carbohydrate, despite the fact that two NXS sequences are present. It appears that all of fibrinogen's carbohydrate chains are of biantennary structure. Variable desialylation, or removal of the N-acetylneuraminic acid, accounts for part of the heterogeneity of circulating fibrinogen. Fibrinogen isolated from human plasma contains equal amounts of mono- and disialyated chains, but no asialo chains (Townsend *et al.*, 1982, 1984).

The carbohydrate on fibrinogen has striking consequences for fibrin polymerization and clot structure. Patients with cirrhosis of the liver and some other liver diseases have fibrinogen with high levels of sialyation of their carbohydrate, leading to clots made up of thin fibers with many branch points (Martinez *et al.*, 1983). These results are consistent with studies using neuraminidase to remove the sialic acid from the carbohydrate of normal fibrinogen, producing clots made up of thicker fibers (Dang *et al.*, 1989). Complete removal of carbohydrate has more dramatic effects on clot structure, resulting in clots made up of massive fibers (Langer *et al.*, 1988). These results suggest that both charge and mass of the carbohydrate help to modulate the extent of lateral aggregation, and

that the carbohydrate clusters significantly enhance (if not largely account for) fibrinogen's solubility per se.

E. X-Ray Crystallographic Structure

Crystallization of large, flexible, heterogeneous proteins like fibrinogen is difficult. Bovine fibrinogen was first crystallized after proteolytic removal of C-terminal portions of the Aα chains with a bacterial enzyme possessing unique specificity (Tooney and Cohen, 1972, 1974). Both crystals and microcrystals of this modified fibrinogen were examined by electron microscopy and analyzed to provide information about fibrinogen structure and the molecular packing in fibrin (Cohen *et al.*, 1983; Weisel *et al.*, 1978, 1981). These crystal forms are unusual in that they are made up of end-to-end bonded molecules that form flexible filaments. The fibrinogen molecule was demonstrated to be about 45 nm in length, containing a central region and two globular regions at each end connected by rods (Weisel *et al.*, 1985). The distal end domain was identified as the C-terminal γ chain, while the C-terminal β chain was folded back near the coiled coil. Cryoelectron microscopy combined with X-ray crystallography yielded a low-resolution structure of fibrinogen (Rao *et al.*, 1991).

The atomic resolution structure of fibrinogen has been accumulated through crystallography of smaller fragments of this protein. The structure of the C-terminal portion of the γ chain was solved, followed shortly by that of fragment D. Subsequently, the structure of a unique fragment E was solved, as well as atomic level structures of modified bovine fibrinogen and chicken fibrinogen (Fig. 4).

The 30 kDa globular C-terminal portion of the human γ chain was expressed in yeast, crystallized, and its structure was determined (Yee *et al.*, 1997). Three domains were identified in the structure, including a C-terminal fibrin-polymerization domain that contains a single calcium-binding site and a deep binding pocket for the complementary knob in the middle of the molecule, exposed by cleavage of the A fibrinopeptide. The C-terminal residues containing two factor XIIIa crosslinking sites and the platelet receptor recognition site were not resolved, probably because they are highly flexible and/or disordered.

The structure of fragment D from human fibrinogen was solved and shown to consist of a coiled-coil region and the two homologous C-terminal $\gamma-$ and $\beta-$nodules oriented at approximately 130 degrees to each other (Spraggon *et al.*, 1997). The structure of the Factor XIIIa crosslinked dimer of fragment D (D dimer or double-D) was solved in the presence and absence of the peptide ligands that simulate the

two knobs exposed by the removal of fibrinopeptides A and B, respectively. The peptide Gly-Pro-Arg-Pro-amide, an analog of the knob exposed by the thrombin-catalyzed removal of fibrinopeptide A, was found to reside in the γC holes; the peptide Gly-His-Arg-Pro-amide, which corresponds to the β-chain knob, was found in the homologous βC holes, demonstrating that the β-chain knob can potentially bind to a homologous hole on the C-terminal βC nodule (Fig. 4; Everse *et al.*, 1998). The γC and βC holes are structurally very similar, but they are able to distinguish between these two peptides that differ by a single amino acid. However, in the absence of the peptide ligand Gly-Pro-Arg-Pro-am, Gly-His-Arg-Pro-am binds to the γC pocket. Additionally, the βC nodule, like its γC counterpart, binds calcium. Comparison of the structures revealed a series of conformational changes brought about by the various knob-hole interactions (Everse *et al.*, 1999; Kostelansky *et al.*, 2002). In particular, a movable ''flap'' of two negatively charged amino acids (Gluβ397 and Aspβ398) has side chains that are pinned back to the coiled coil with a calcium atom bridge until Gly-His-Arg-Pro-am occupies the βC pocket.

The structure of modified bovine fibrinogen provided information about the overall form of the molecule and the chain arrangement in the central region where the two halves of the molecule are joined, as well as the thrombin-binding site (Brown *et al.*, 2000). However, the N-terminal ends of the $\alpha-$ and $\beta-$chains were not visible, likely because they are flexible and/or disordered. The coiled-coil region of fibrinogen has a planar, sigmoidal shape. Near the middle of the coiled-coil region, at the plasmin-sensitive segment, there is a hinge about which the molecule adopts different conformations, which may help accommodate variability in the structure of the fibrin clot. An improved model of intermediates in fibrin polymerization was developed (Brown *et al.*, 2000).

Since chicken fibrinogen has much smaller αC connectors, it was possible to crystallize it so that its structure could be solved (Yang *et al.*, 2001). The structure included the sigmoidally shaped coiled coil, as well as the planar disulfide rings in the middle. Both the αC domains and the N-terminal segments of the Aα- and Bβ-chains, including fibrinopeptides A and B, were disordered and/or flexible. The crystal structure of chicken fibrinogen complexed with two synthetic peptides was determined. In the coiled-coil regions, there is a strong correlation between mobility of residues and plasmin attack sites.

The crystal structure of the central region of bovine fibrinogen, a fragment E, provided a view of the central region of the molecule and proximal portions of the coiled coil (Madrazo *et al.*, 2001). The two halves of the molecule are joined together at the center in an extensive

molecular ''handshake'' by using both disulfide linkages and noncovalent contacts. On one face of the fragment, the Aα− and Bβ−chains from the two half molecules form a funnel-shaped domain with a hydrophobic cavity. On the opposite face, the N-terminal γ chains fold into a separate domain. Despite the chemical identity of the two halves of fibrinogen, an unusual pair of adjacent disulfide bonds locally constrain the two γ chains to adopt different conformations. The striking asymmetry of this domain may promote the known twisting of the intermediates in fibrin polymerization (Weisel *et al.*, 1987).

The crystal structures of fragment D from several recombinant fibrinogens has demonstrated that these mutant molecules are excellent models for the study of structure-function relationships, since the polypeptide chains are folded properly with only local changes (Kostelansky *et al.*, 2002). Moreover, these studies allow testing of hypotheses about the effects of a particular mutation on structure and internal interactions within the molecule, which then can be correlated with the effects on fibrin polymerization or platelet aggregation (Kostelansky *et al.*, 2004a,b).

F. Calcium Binding

Fibrinogen has both strong and weak binding sites for calcium ions, which are important for its structural stability and functions. High-affinity binding sites for calcium are present in the γ chains, with residues Asp318, Asp320, Gly324, Phe322, and two strongly bound water molecules providing coordination in sites with amino acid sequence homologous to the EF hand calcium binding site in calmodulin (Spraggon *et al.*, 1997; Yee *et al.*, 1997). However, the calcium binding loop γ311-336 is not a classic EF hand, but a unique structure. The dissociation constant for calcium at these binding regions is such that these sites will be fully occupied in plasma fibrinogen. Although it was thought that there is a high-affinity calcium-binding site in the central region, none has yet been visualized.

Some lower-affinity sites are associated with interactions of the C-terminal Bβ chain with the coiled-coil backbone (Everse *et al.*, 1999; Kostelansky *et al.*, 2002). Other low-affinity binding sites are less well defined. There is evidence that some of these sites may be associated with the sialic acid residues on the carbohydrate chains (Dang *et al.*, 1989).

When calcium ions occupy the γ-chain high-affinity sites, the γ chains are protected from enzymatic degradation (Ly and Godal, 1973; Odrljin *et al.*, 1996). The peptide Gly-Pro-Arg-Pro protects the molecule from plasmin digestion in a similar manner (Yamazumi and Doolittle, 1992).

Calcium binding does not affect fibrinopeptide A cleavage by thrombin or batroxobin, but appears to be important for modulating fibrin polymerization by enhancing lateral aggregation to form thicker fibers. There is a conformational change in fibrin associated with fibrinopeptide B cleavage that is calcium-dependent (Donovan and Mihalyi, 1985; Dyr *et al.*, 1989; Mihalyi, 1988a). There is a change in affinity of calcium-binding sites associated with fibrinopeptide B release and this conformational change. Gly-His-Arg-Pro binds tenfold more strongly to fibrinogen in the presence of calcium than in its absence. This may be related to the low-affinity binding site for calcium in the βC nodule that is involved in the conformational change associated with binding of Gly-His-Arg-Pro (Everse *et al.*, 1999).

III. Biosynthesis and Metabolism of Fibrinogen

Human fibrinogen is the product of three closely linked genes, each specifying the primary structure of one of the three polypeptide chains (Chung *et al.*, 1983, 1990; Crabtree, 1987). The ability to express human fibrinogen in mammalian cells provides a means to study the synthesis and secretion of fibrinogen. Site-directed mutagenesis of fibrinogen has established the sequence of steps in fibrinogen assembly. There is a progression from single chains to two-chain complexes to trimeric half molecules, which then dimerize to form a whole fibrinogen molecule. The fibrinogen genes are located in a cluster in the distal third of chromosome 4, bands q23-132 (Kant *et al.*, 1985). The gene for the fibrinogen Aα chain, located in the middle of the gene cluster, consists of five exons that code for 625 amino acids, but 15 amino acids are removed posttranslationally. An extension of an open reading frame into an alternately spliced sixth intron gives rise to a longer Aα chain (αE) in 1–2% of molecules in the adult (more in fetal blood), so that they contain an additional domain homologous to the C-terminal Bβ and γ chains, leading to a fibrinogen molecule with a molecular mass of 420 kDa (Fu and Grieninger, 1994). The gene for the Bβ chain, containing eight exons, is located downstream to that of the Aα and γ chains and is transcribed in the opposite direction. The γ chain gene consists of ten exons, but alternative processing of the γ chain transcript leads to read-through of the exon IX/intron I splice junction, which produces 10–15% of plasma fibrinogen molecules with the γ' chain, containing a 20-amino acid extension at the C-terminus (Chung and Davie, 1984; Wolfenstein and Mosesson, 1981).

Fibrinogen synthesis is stimulated by hormones such as the glucocorticoid dexamethasone or suppressed by the estrogen estadiol-17 (Amrani *et al.*, 1983; Wangh *et al.*, 1983). The three fibrinogen genes share a

common regulatory mechanism, as indicated by the highly coordinated expression of mRNAs, such that in hepatocytes the relative proportions of the mRNAs for the three chains are nearly the same (Crabtree, 1987). The regulation may involve similar sequence motifs in the immediate 5′ flanking regions and common transcriptional regulatory molecules (Hu *et al.*, 1995). In the absence of serum, the *in vitro* expression of the Aα chain is uncoordinated with expression of the Bβ and γ chains, suggesting that signal transduction factors in serum play a part in regulation (Plant and Grieninger, 1986). Some of the general and liver-specific transcription factors involved in the regulation of fibrinogen expression have been identified.

Fibrinogen is one of the proteins upregulated during the acute phase response, with increases two- to tenfold, while albumin levels decrease by about 50% (Crabtree, 1987). Under stressful conditions, there are a series of reactions that result in cell activation and cytokine production. The cytokines then act on other cells, resulting in an increase in production of glucocorticoids, proliferation of immune cells, and changes in the levels of plasma proteins synthesized in the liver. The upregulation of the acute-phase response proteins is mediated by interleukin-6 (IL-6) (Hantgan *et al.*, 2000). IL-6 response elements are located on all three fibrinogen genes in the 5′ flanking regions. Glucocorticoid response regions are present in the 5′ flanking region of the Bβ and γ genes.

Fibrinogen is assembled in the rough endoplasmic reticulum of the liver. In human hepatocellular carcinoma cells, the newly synthesized Bβ chains are secreted and used more rapidly than the Aα and γ chains, so that a large intracellular pool of free Aα and γ chains exists (Yu *et al.*, 1984). In contrast, the Aα chains are the limiting factor in cultured rat or chicken hepatocytes (Hirose *et al.*, 1988; Plant and Grieninger, 1986). Stimulation by serum results in a 20-fold increase in secretion of fibrinogen and comparable levels of all three chains (Baumhueter *et al.*, 1989).

The assembly of polypeptide chains into functional fibrinogen molecules has been studied in several expression systems. The order in which the polypeptide chains are joined together by disulfide bonds has been determined and specific structural features that are important for assembly, such as the coiled-coil, have been identified (Xu *et al.*, 1996). Substitution of the cysteines with serine showed that the interchain disulfide ring at the proximal end of the coiled-coil, in addition to the disulfides between the two halves of the molecule, are necessary for assembly of the two half molecules (Zhang and Redman, 1996). The distal interchain disulfide ring is not necessary for assembly of the two half molecules but is necessary for secretion. Disruption of intrachain disulfide bonds and deletions of portions of the polypeptide chains have revealed which

disulfide bonds and noncovalent interactions are necessary for fibrinogen assembly.

The liver is the primary source of plasma fibrinogen, with a steady state rate of synthesis of 1.7–5.0 g per day and a large reserve (Takeda, 1966). Three-quarters of the human fibrinogen is located in the plasma but it is also present in platelets, lymph nodes, and interstitial fluid. The half-life of fibrinogen is 3 to 5 days, but the catabolic pathway is different from the classic plasma protein pathways and is largely unknown. Coagulation and lysis accounts for only 2–3% of fibrinogen loss in healthy individuals (Nossel, 1976). Fibrinogen degradation products appear to play a role in regulation of the fibrinogen turnover (Nham and Fuller, 1986).

Platelets contain fibrinogen but there is controversy over its origin, and whether it is structurally and functionally distinct from plasma fibrinogen. The γ chain variants of fibrinogen do not appear to be present in platelets (Haidaris *et al.*, 1989); in one patient with a heterozygous dysfibrinogen, the platelets contained only normal fibrinogen (Jandrot-Perrus *et al.*, 1979). The megakaryocytes, from which platelets originate, are a possible source of fibrinogen synthesis. However, in subjects with hypofibrinogenemia, there are also lower levels of fibrinogen in platelets; infusion of fibrinogen results in a subsequent increase in platelet fibrinogen (Harrison *et al.*, 1989). Also, no fibrinogen mRNA was detected in megakaryocytes (Louache *et al.*, 1991). Therefore, it seems that platelet and megakaryocyte fibrinogen arise primarily from αIIbβIII-mediated endocytosis of plasma fibrinogen and its storage in α granules.

Fibrinogen is also synthesized in several extrahepatic sites, suggesting that it may function in other structural or adhesive capacities, but there is little evidence for this. Fibrinogen γ chain gene expression has been shown *in vivo* only in bone marrow, brain, and lung (Haidaris and Courtney, 1990). Epithelial cells from lung and intestine secrete small amounts of fibrinogen in a polarized manner from their basolateral face (Haidaris, 1997). It may be that the lung epithelium secretes fibrinogen and incorporates it into the extracellular matrix under certain pathological conditions, contributing to fibrotic lung disease. Synthesis of fibrinogen by cultured granulosa cells may reflect a possible function for it in ovulation (Parrott *et al.*, 1993). The apparent *in vivo* synthesis of fibrinogen by trophoblasts (Galanakis *et al.*, 1996) and the fact that the trophoblast basement membrane consists largely of fibrin(ogen) suggest that these cells may secrete fibrinogen into their abluminal and/or interstitial environment as their need dictates, but the functional significance remains an open question.

IV. Conversion of Fibrinogen to Fibrin

Fibrin polymerization is initiated by the enzymatic cleavage of the fibrinopeptides, converting fibrinogen to fibrin monomer (Fig. 1). Then, several nonenzymatic reactions yield an orderly sequence of macromolecular assembly steps. Several other plasma proteins bind specifically to the resulting fibrin network. The clot is stabilized by covalent ligation or crosslinking of specific amino acids by a transglutaminase, Factor XIIIa.

A. *Thrombin-Catalyzed Cleavage of Fibrinopeptides*

Thrombin, which is produced on proteolytic cleavage of prothrombin by Factor Xa in the presence of Factor V and phospholipid, is a serine protease with specificity for fibrinogen's fibrinopeptides at particular Arg-Gly bonds. This specificity arises in part because of hydrophobic and structure-dependent interactions between the fibrinopeptides and thrombin's catalytic site, as well as noncatalytic binding of enzyme to substrate (Stubbs and Bode, 1993). Distinct regions of thrombin are involved in catalysis and fibrin binding. In addition to binding at the catalytic site, thrombin also binds to the central region of fibrinogen through ionic interactions with positively charged residues on the thrombin B chain, termed the *anion-binding exosite I*. Evidence for this binding site first arose from the observation that active thrombin, as well as inactivated forms, bind to fibrin. Solution of the structure of thrombin showed the details of the catalytic site and the exosites (Bode and Stubbs, 1993).

Studies of the salt and temperature dependence of fibrinopeptide release show that electrostatic interactions allow thrombin to bind at diffusion-controlled rates (De Cristofaro and Di Cera, 1992; Vindigni and Di Cera, 1996). After thrombin is bound, hydrophobic interactions result in a conformational change in this allosteric enzyme, converting it to a faster form with a higher catalytic efficiency (Guinto *et al.*, 1995).

The crystal structure of a complex between thrombin and the E fragment corresponding to the central region of fibrinogen shows how the two molecules interact (Pechik *et al.*, 2004). The complex consists of two thrombin molecules bound to opposite sides of the central part of fragment E, such that there is the correct orientation of their catalytic triads for cleavage of the fibrinopeptides of fibrinogen. As expected, binding occurs through thrombin's anion-binding exosite I, but only part of it is involved in forming an interface with the complementary negatively charged surface of fragment E.

A fibrin precursor, with one A fibrinopeptide removed and one intact, has been isolated (Shainoff *et al.*, 2001). Although fibrinopeptide B release begins at the same time as that of the A fibrinopeptides, the experimental data are fit by a sequential model in which the B fibrinopeptides are released at a significant rate only after A fibrinopeptide cleavage (Lewis *et al.*, 1985; Mihalyi, 1988b). With polymerization of fibrin, the rate of B fibrinopeptide release increases about sevenfold, which indicates that cleavage occurs preferentially from fibrin polymers or that there is a polymerization-induced conformational change that enhances release (Hanna *et al.*, 1984; Martinelli and Scheraga, 1980).

Substitution of particular amino acids in the fibrinopeptides and at thrombin's cleavage sites in recombinant fibrinogen has been used to understand thrombin specificity, the relative rates of fibrinopeptide cleavage, and effects on polymerization. If fibrinopeptide B is replaced with a fibrinopeptide-A-like peptide, fibrinopeptide release is similar to release from the Aα chain (Mullin *et al.*, 2000). Thus, the ordered release of fibrinopeptides is dictated by the specificity of thrombin for its substrates.

Thrombin binds strongly to fibrin in the clot or thrombus so that thrombin is sequestered and is less likely to be inactivated or prolong clotting (Mosesson, 2003). Release and exchange of bound thrombin for that in solution, as a result of fibrinolysis or rinsing, is much slower than thrombin adsorption. Fibrin has two classes of thrombin binding sites, differing in affinity. The high-affinity site has been shown to be on the γ' chain of fibrinogen (Lovely *et al.*, 2003; Meh *et al.*, 2001). Studies of dysfibrinogenemias with mutations, such that clots do not bind thrombin as strongly, demonstrate the clinical significance of fibrin's antithrombin activity, in that these patients tend to have serious thromboembolic disease.

B. Polymerization Steps

Although the fibrinopeptides constitute less than 2% of the mass of the fibrinogen molecule, their cleavage has profound consequences for the solubility of the resulting fibrin. Cleavage of the A fibrinopeptides exposes binding sites in the central domain (called *A*) that are complementary to sites (called *a*) always exposed at the ends of the molecules (Fig. 5) (Budzynski, 1986; Doolittle, 1984). The *A* ''knobs'' consist in part of the newly exposed N-terminus of fibrin's α chain, Gly-Pro-Arg, but also include part of the Bβ chain. The *a* sites include ''holes'' in the C-terminal γ chain (Fig. 4). There are also *B* ''knobs'' exposed in the middle of the molecule on cleavage of the B fibrinopeptides after polymerization begins; these

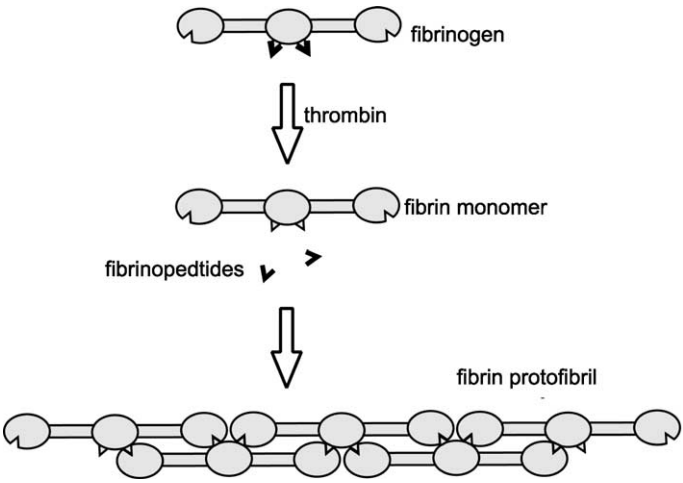


FIG. 5. Schematic diagram of complementary binding sites or knob-hole interactions in fibrin polymerization. Fibrinopeptides in the central domain cover knobs that are complementary to holes that are always exposed at the ends of the protein. When the fibrinopeptides are removed by thrombin, knob-hole interactions occur, giving rise to the two-stranded protofibril made up of half-staggered molecules.

sites are complementary to *b* ''holes'' in the C-terminal β chain (Fig. 4) (Medved *et al.*, 1993; Shainoff and Dardik, 1983).

Specific interactions between the *A:a* complementary binding sites produce aggregates in which the fibrin monomers are half-staggered, since the central domain of one molecule binds to the end of the adjacent molecule (Fig. 5). Initially, a dimer is formed and then additional molecules are added to give a structure called the *two-stranded protofibril* (Fig. 3B) (Fowler *et al.*, 1981; Medved *et al.*, 1990).

Once protofibrils reach a sufficient length (usually about 600–800 nm), they aggregate laterally to form fibers (Fig. 6) (Hantgan and Hermans, 1979; Hantgan *et al.*, 1980). The intermolecular interactions that occur in lateral aggregation are specific so that the fibers have a repeat of 22.5 nm, or about half the molecular length, and a distinctive band pattern as observed by electron microscopy of negatively contrasted specimens (Fig. 3C) (Cohen *et al.*, 1963; Weisel, 1986a) or X-ray fiber diffraction (Stryer *et al.*, 1963). The band pattern directly reflects the molecular structure and packing in fibrin and indicates that fibers are paracrystalline structures with the molecules precisely aligned in the longitudinal direction, but only partly ordered in the lateral direction (Freyssinet *et al.*, 1983; Weisel, 1986a; Weisel *et al.*, 1983).

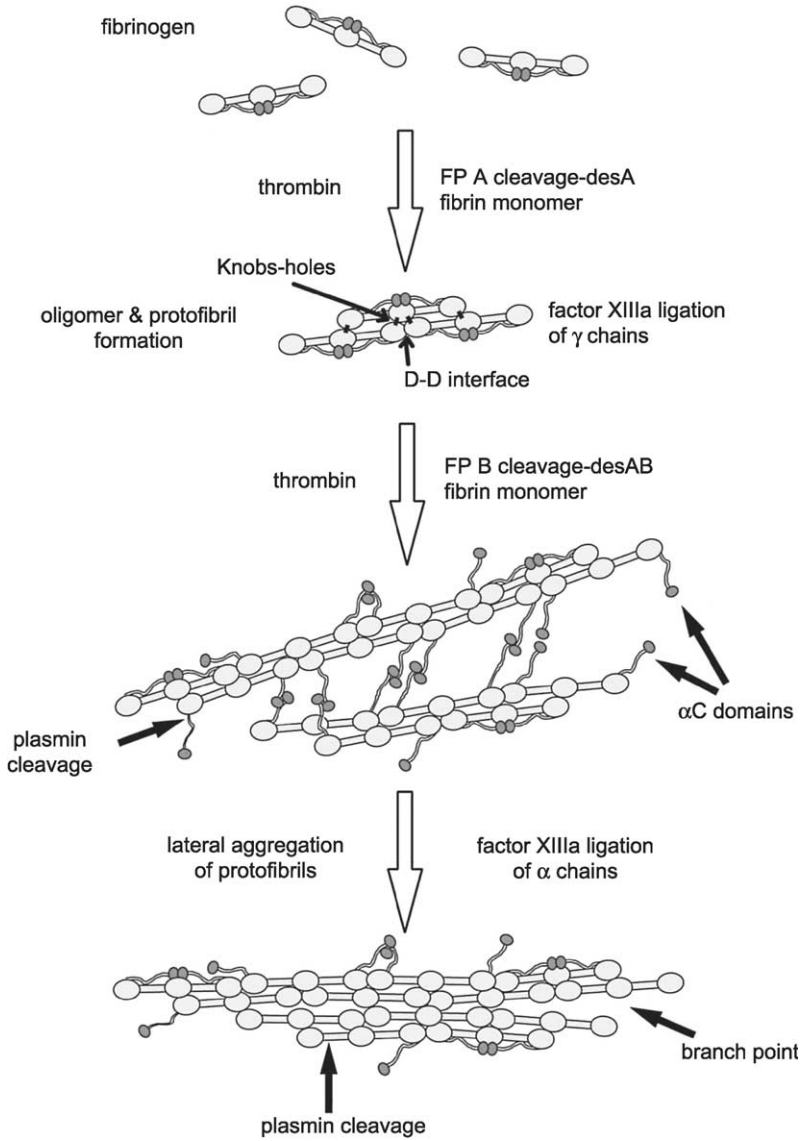


FIG. 6. Schematic diagram of fibrin polymerization. Fibrinopeptide A is cleaved from fibrinogen, producing desA fibrin monomers, which aggregate via knob-hole interactions to make oligomers. Fibrinopeptide B is cleaved primarily from polymeric structures. The oligomers elongate to yield protofibrils, which aggregate laterally to make fibers, a process enhanced by interactions of the αC domains. Factor XIIIa crosslinks or ligates γ chains more rapidly than α chains. Plasmin cleaves the αC domains and Bβ1-42 and then cuts across the fibrin in the middle of the coiled coil. At the bottom of the diagram, a branch point has been initiated by the divergence of two protofibrils.

Snake venom enzymes that remove only the A fibrinopeptides have been used to show that cleavage of the B fibrinopeptides enhances lateral aggregation, producing thicker fibers (Blombäck *et al.*, 1978). Other snake venom proteases that preferentially cleave the B fibrinopeptides yield clots only at low temperature (Shainoff and Dardik, 1979). Clots formed via cleavage of either only fibrinopeptide A or only fibrinopeptide B are made up of fibers with normal appearance, indicating that both the *B* knobs and the *A* knobs can participate in protofibril formation and neither are specifically associated with lateral aggregation (Mosesson *et al.*, 1987; Weisel, 1986b). Studies with recombinant mutant fibrinogen have shown that an intact *a* site is necessary for protofibril formation with cleavage of the B fibrinopeptides, suggesting that *B:a* interactions may be involved (Hogan *et al.*, 2001). There does not appear to be any clear evidence for *B:b* interactions in fibrin.

The αC domains enhance lateral aggregation in fibrin polymerization and are important for the mechanical properties and stability of clots (Fig. 6) (Weisel and Medved, 2001). The αC domains can be seen as extensions or appendages of some fibrinogen molecules (Erickson and Fowler, 1983; Mosesson *et al.*, 1981; Price *et al.*, 1981; Weisel *et al.*, 1985), but normally these regions are interacting with the central portion of the molecule (Veklich *et al.*, 1993). On the cleavage of the B fibrinopeptides, there is a large-scale conformational change such that the αC domains dissociate from the central region and are available for intermolecular interactions (Veklich *et al.*, 1993). Experiments with purified or recombinant preparations missing either one or both of the αC domains, αC fragments, and several dysfibrinogenemias indicate that intermolecular interactions between αC domains are important for the enhancement of lateral aggregation during fibrin polymerization (Gorkun *et al.*, 1994; Veklich *et al.*, 1993; Weisel and Medved, 2001).

Lateral growth of fibers seems to be limited because protofibrils in the fibers are twisted (Fig. 7A), so that as the fiber diameter increases, they must be stretched to traverse an increasingly greater path length (Weisel *et al.*, 1987). With this model, fiber growth stops when the energy required to stretch an added protofibril exceeds the energy of bonding.

The fibrils making up the fibers branch, leading to a three-dimensional network (Fig. 7A). This property of branching is essential for the properties of fibrin since it leads to the production of a space-filling gel. Such a gel can be formed even with very low fibrinogen concentrations (<0.01 mg/ml protein), and clots made from 50 mg/ml fibrinogen are still 95% liquid. Branching can be initiated by a fibrin monomer that binds to one other monomer at its end, and then diverges from that molecule to interact with a second monomer in its central region, producing a

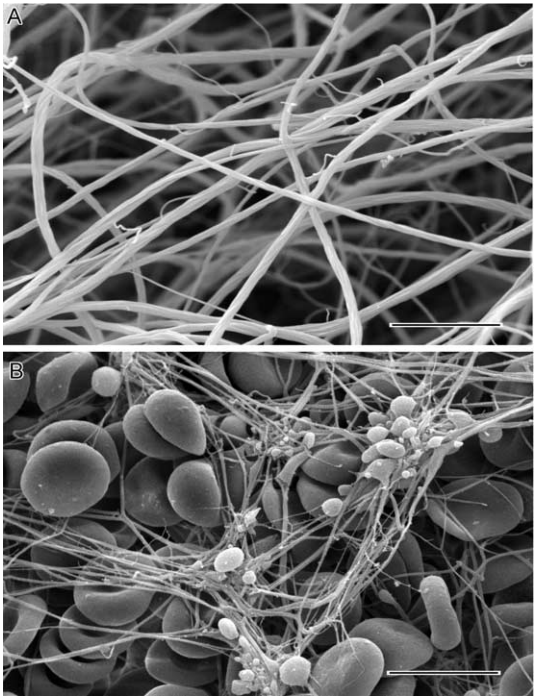


FIG. 7. Scanning electron microscope images of fibrin clots. (A) Electron micrograph of clot formed by addition of thrombin to purified fibrinogen. A clot network is formed by branching of fibers, with trifunctional junctions, but in these images it is sometimes difficult to distinguish between true branch points and fibers crossing at different depths in the clot. Note that the twisting of protofibrils making up the fibers is visible on the surface of some fibers. Magnification bar, 5 μm. (Image from Weisel, 2004b.) (B) Electron micrograph of whole blood clot, made from freshly drawn blood with no additions. Fibrin fibers commonly originate from platelet aggregates and erythrocytes and leukocytes are trapped in the meshwork. Magnification bar, 10 μm. (Image from Yuri Veklich and John Weisel, University of Pennsylvania.)

trimolecular branch point (Mosesson *et al.*, 1993). In addition, other branch points are simply points at which two parallel strands diverge from each other (Fig. 6) (Baradet *et al.*, 1995; Hantgan and Hermans, 1979). In any case, electron microscopy has revealed that branch points are almost invariably trifunctional and are generally made up of fibers of similar diameter diverging at a small angle (Fig. 3C) (Ryan *et al.*, 1999; Weisel, 2004b). The physical and mechanical characteristics of clot networks vary greatly depending on the conditions of polymerization.

Interactions between fibrin and fibrinogen are likely to be important during early stages of polymerization, when fibrin monomers are

present in small amounts and will exist primarily as soluble complexes with fibrinogen (Sasaki *et al.*, 1966). Similarly, such complexes are likely to affect polymerization of clots made with low thrombin, such as by addition of tissue factor to whole blood in the presence of contact phase inhibitors (Rand *et al.*, 1996). In addition, thrombus growth *in vivo* could be affected by fibrinogen or fibrin degradation products by competing for polymerization sites.

Studies of plasma clots made under a variety of conditions suggest that the network is established early as a result of the activation process (Blombäck *et al.*, 1994) and that the thrombin activation pathway modulates clot structure (Torbet, 1995). Light scattering studies, combined with other physical chemical techniques, have substantiated this conclusion for purified fibrinogen and provided more information about the structure and properties of the assembling clot (Bernocco *et al.*, 2000; Ferri *et al.*, 2001, 2002; Profumo *et al.*, 2003).

C. *Binding Pockets*

X-ray crystallography studies have defined the structure of the *a* and *b* binding pockets or holes in the C-terminal portions of the γ and β chains, respectively (Fig. 4). These binding pockets are always exposed, allowing even fibrinogen-fibrin interactions, as just mentioned. The *a* holes bind synthetic Gly-Pro-Arg-Pro peptides, while the *b* holes preferentially bind Gly-His-Arg-Pro peptides. However, depending on the conditions, such synthetic peptides may bind to the alternative binding pocket.

The interactions of Gly-Pro-Arg-Pro with residues in the *a* pocket define minimum features of the binding, but the interaction site may be more extensive. Gly-Pro-Arg-Pro is held in the hole mainly by electrostatic interactions with residues including Gln329, Asp330, His340, and Asp364 (Spraggon *et al.*, 1997). In contrast to the small structural changes that accompany Gly-Pro-Arg-Pro binding to the *a* hole, a major conformational change accompanies binding of Gly-His-Arg-Pro to the *b* hole (Everse *et al.*, 1999; Kostelansky *et al.*, 2002). The βC nodule, which is interacting with the coiled-coil calcium bridge, moves away and two carboxylate anions rotate inward to complete formation of the pocket, with accompanying release of calcium.

D. *Other Binding Interactions*

The half-staggered interactions present in fibrin also create a D–D contact at the end-to-end junction of two fibrin molecules (Fig. 6). This D–D interface is an important feature in most of the crystal forms of

fibrinogen and its fragments, so it has been well characterized (Brown *et al.*, 2000; Spraggon *et al.*, 1997; Weisel *et al.*, 1978). At this interface, γArg275 makes contacts with both γTyr280 and γSer300; γ308Asn and other residues are also involved in these interactions.

There is evidence for the involvement of the N-terminal Bβ chain in fibrin polymerization (Pandya *et al.*, 1985; Siebenlist *et al.*, 1990). A recombinant fibrinogen with His substituted for Arg at the Bβ thrombin-cleavage site led to a 300-fold decrease in the rate of fibrinopeptide B release, whereas the rate of fibrinopeptide A release was normal (Moen *et al.*, 2003). As a consequence, thrombin- or batroxobin-catalyzed or desA monomer polymerization was impaired, due to the histidine substitution itself. Thus, it appears that the N-terminus of the Bβ chain is involved in the lateral aggregation of normal desA protofibrils.

E. Covalent Crosslinking of Fibrin by Factor XIIIa

The clot is stabilized by the formation of covalent bonds introduced by a plasma transglutaminase, Factor XIIIa (Greenberg *et al.*, 2003; Henschen and McDonagh, 1986; Loewy *et al.*, 2000; Lorand, 2001), which is necessary to make fibrin resistant to mechanical and proteolytic insults to prevent bleeding. Factor XIII is a 326 kDa tetrameric complex with two A chains and two B chains. In addition, about half of the total Factor XIII in blood arises from platelets, which contain dimers consisting of only two A chains. The three-dimensional structure of the recombinant A chain of human Factor XIII was determined by X-ray crystallography, showing that each A chain is folded into four domains: catalytic core, β sandwich, barrel 1, and barrel 2 (Yee *et al.*, 1994).

The active enzyme, Factor XIIIa, is generated from its precursor, Factor XIII, by the action of thrombin in the presence of calcium and fibrin (Fig. 1). The 37-amino-acid activation peptide prevents substrates from entering the active site pocket containing Cys 314, with the activation peptide from one A chain blocking the active site of the other A chain. Activation of Factor XIII involves thrombin cleavage of the A_2B_2 zymogen form of Factor XIII, followed by calcium-dependent dissociation of the A_2 subunits (Greenberg *et al.*, 2003; Lorand, 2001). In plasma, fibrinopeptide A is cleaved more than 40 times faster than the Factor XIII activation (Greenberg *et al.*, 1985). Factor XIII binds strongly via its B chain to fibrinogen, so that nearly all circulating zymogen is bound (Greenberg *et al.*, 2003). The primary site of binding of Factor XIII to fibrinogen appears to be the γ′ splice variant (Moaddel *et al.*, 2000a,b; Siebenlist *et al.*, 1996). This interaction plays a role in maintaining an association between these two proteins from before clot formation.

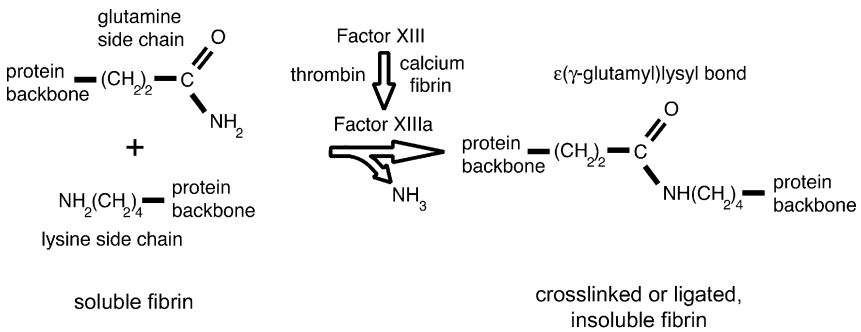


FIG. 8. Formation of isopeptide bond catalyzed by Factor XIIIa. The chemical reaction was catalyzed by Factor XIIIa, yielding insoluble fibrin crosslinked by Nε-(γ glutamyl) lysine bonds. Factor XIII is activated to Factor XIIIa by thrombin in the presence of calcium ions and fibrin.

Fibrin polymers are responsible for the fibrin-dependent enhancement of Factor XIII activation (Greenberg *et al.*, 2003). The mechanism for this effect involves the formation of a tight ternary complex between fibrin, Factor XIII, and thrombin, accompanied by a conformational change of Factor XIII that exposes the active site, after which Factor XIIIa remains bound to fibrin. However, the B chains dissociate, which is necessary to expose the active site cysteine of plasma Factor XIII. Platelet Factor XIII without the B chains, is more rapidly activated by thrombin than plasma Factor XIII because of the time that it takes for the B chains to dissociate.

Many isopeptide bonds can be formed between the side chains of ε-lysine (donor) and γ-glutamine (acceptor) residues (Fig. 8). The γ chains of fibrinogen are crosslinked or ligated first, followed by the C-terminal α chains. In addition, other proteins—notably α2-antiplasmin, plasminogen activator inhibitor-2, and fibronectin—are also covalently ligated to fibrin by Factor XIIIa (Greenberg *et al.*, 2003).

The γ chain residues involved in Factor XIIIa-induced crosslinking, γGln398/399 and γLys406, are located near the C-terminal end of the γ chain (Greenberg *et al.*, 2003). Factor XIIIa catalyzes the formation of crosslinks between the γ chains of fibrin at a much faster rate than it does to those of fibrinogen. It has been demonstrated by using a synthetic construct with two spaced Gly-Pro-Arg-Pro peptides that this effect is a result of physical proximity of the acceptor and donor sites in fibrin (Lorand *et al.*, 1998). The crystal structure of crosslinked human D dimer does not show the end of the γ chains where the crosslinks are present because of the flexibility of this part of the molecule, but half of one crosslink is present in the structure of lamprey D dimer (Yang *et al.*, 2002).

There has been an ongoing controversy on whether the γ chains of fibrin are crosslinked longitudinally or transversely, and the evidence for both viewpoints has been summarized elsewhere (Mosesson, 2004a,b; Weisel, 2004a,c).

Each α chain contains potential glutamine acceptor sites at 221, 237, 328, and 366, and donor sites at lysine 508, 539, 556, 580, and 601 (Greenberg *et al.*, 2003; Matsuka *et al.*, 1996). Since the αC domains associate even in the absence of crosslinking, these interactions probably bring acceptor and donor sites in proximity, facilitating the formation of the isopeptide bonds. These bonds create a covalently connected network of αC domains, although little is known of its structure. In addition, there are lesser amounts of γ trimers, tetramers, and α-γ complexes. Factor XIII polymorphisms can have effects on the structure and properties of the fibrin clot (Ariëns *et al.*, 2002).

F. Mechanical Properties of Clots

The mechanical properties of fibrin are essential for its functions in hemostasis and wound healing, since the clot must stop bleeding and yet allow the penetration of cells. The mechanical properties of a thrombus will determine how it responds to flowing blood, including elastic or plastic deformation and embolization. Epidemiological studies have revealed a relationship between myocardial infarction and clot stiffness (Fatah *et al.*, 1996; Scrutton *et al.*, 1994).

Many factors that affect clot structure have a great impact on fibrin's mechanical properties. The thrombin concentration can have an effect on the clot structure, especially if the rate of generation of fibrin monomer by cleavage of fibrinopeptides becomes limiting relative to the rate of polymerization (Blombäck *et al.*, 1994; Weisel and Nagaswami, 1992). In fact, it appears that many factors (including pH, calcium and chloride ion concentration, and other plasma proteins) exert an influence on clot structure and mechanical properties by affecting the rates of various steps in the polymerization process (Carr and Hardin, 1987; Carr *et al.*, 1985, 1986; Di Stasio *et al.*, 1998; Weisel and Nagaswami, 1992).

Fibrin is a viscoelastic polymer, which means that it has both elastic and viscous properties (Ferry, 1988). Thus, the properties of fibrin may be characterized by stiffness or storage modulus (representing its elastic properties) and creep compliance or loss modulus/loss tangent (representing its inelastic properties). These parameters will determine how the clot responds to the forces applied to it in flowing blood. For example, a stiff clot will not deform as much as a less stiff one with applied stress.

A clot with a greater inelastic component will deform permanently with stress, while one with a greater elastic component will return to its original shape.

The mechanical properties have been measured for many different types of clots under various conditions (Ferry, 1988). Both elastic and inelastic properties are very sensitive to small changes that affect polymerization and clot structure. Clots ligated with Factor XIIIa tend to have a higher stiffness and a lower inelastic component of deformation. At low strains, stress is directly proportional to strain so that the stiffness is constant, but at large strains the stiffness of the clot increases 20-fold (Janmey *et al.*, 1983). This strain hardening or stiffening may be important biologically because it allows fibrin clots to be compliant at low strains and then become stiffer at higher strains so that they are not damaged. The mechanical properties are greatly affected by inclusion of other proteins from plasma and by the effects of platelets and other cells.

The origin of clot elasticity is unknown, although it cannot be rubber-like because the clot has very different structural and mechanical properties (Ferry, 1988). Comparisons of the mechanical properties of clots made under a variety of conditions correlated with the clot structures demonstrate empirical relationships that provide some clues to this long-standing puzzle (Ryan *et al.*, 1999). The most likely explanation for the inelastic properties of fibrin is the slippage of protofibrils past each other, but there may be some role of breakage and reformation of noncovalent bonds (Weisel, 2004b).

V. Binding of Other Proteins to Fibrin(ogen)

A. *Fibronectin and Albumin*

Some proteins, such as plasma fibronectin and albumin, interact with fibrin to alter clot structure and properties, although the former becomes crosslinked to fibrin while the latter does not. As a result of these and other interactions, fibrin clots formed in plasma have very different properties than those made with purified proteins (Blombäck *et al.*, 1994; Carr, 1988; Shah *et al.*, 1987). Albumin has significant effects on the extent of lateral aggregation, yielding either thicker or thinner fibers depending on its concentration and other experimental conditions (Galanakis *et al.*, 1987; Torbet, 1986).

Fibronectin is a large glycoprotein normally present in blood and connective tissue that is made up of a number of different domains serving as binding sites to various cells via specific integrin receptors

and to a number of proteins, including fibrinogen, collagen, and gelatin (Ruoslahti, 1988). Fibronectin is important in cell adhesion and may affect wound healing. The binding region for fibronectin on fibrinogen has been identified (Makogonenko *et al.*, 2002). Fibronectin has distinctive effects on clot structure, with fiber size and density affected by the concentration of fibronectin and whether or not it is covalently linked to fibrin with Factor XIIIa.

B. Thrombospondin, von Willebrand Factor, and Fibulin

Thrombospondin is a 540 kDa glycoprotein released from the α granules of stimulated platelets that has a variety of functions in coagulation, cell adhesion, angiogenesis, and inflammation. Thrombospondin appears to play a role in platelet aggregation by stabilizing the interactions between fibrinogen and platelet integrins that are reversible initially (Bacon-Baguley *et al.*, 1990; Leung, 1984). Thrombospondin is also incorporated into fibrin clots and affects clot structure and properties. Fibrinogen and thrombospondin form a reversible, noncovalent complex. Fragments of thrombospondin containing the interchain disulfide bonds retain the ability to bind to fibrinogen (Dixit *et al.*, 1984). The corresponding region of fibrinogen that binds to thrombospondin appears to reside in either the αC domains (Tuszynski *et al.*, 1985) or in the middle of the Bβ and γ chains (Bacon-Baguley *et al.*, 1990).

Von Willebrand factor is a large glycoprotein that is present in blood as multimers and is important in initiating the adhesion of platelets to sites of vascular damage at the beginning of formation of a platelet plug. Von Willebrand factor can bind to fibrin noncovalently and then can be cross-linked to fibrin with Factor XIIIa (Beguin and Kumar, 1997; Hada *et al.*, 1986). These interactions may be important in the formation of a platelet fibrin thrombus and in anchoring platelets to fibrin in flowing blood.

Fibulin-1 is an extracellular matrix protein that is present in blood and was first isolated in a complex with fibrinogen, to which it binds with high affinity via the C-terminal end of the Bβ chain (Tran *et al.*, 1995). Fibulin-1 can support platelet adhesion to extracellular matrix under flow conditions via a fibrinogen bridge (Godyna *et al.*, 1996).

C. Fibroblast Growth Factor-2, Vascular Endothelial Growth Factor, and Interleukin-1

Fibroblast growth factor-2 binds to fibrin specifically at two classes of binding sites and has striking effects on endothelial cell proliferation (Sahni *et al.*, 1998, 1999). Vascular endothelial growth factor binds

specifically to fibrinogen and fibrin also at two classes of sites (Sahni and Francis, 2000). Interleukin-1β (but not interleukin-1α) binds to fibrinogen and fibrin at the same or similar sites as fibroblast growth factor-2 (Sahni *et al.*, 2004). Since fibrin is part of the provisional matrix needed to support vascular cell responses for repair, the binding of these growth factors and cytokines will localize them to the site of injury so that they can enhance cellular responses needed for wound repair and angiogenesis.

VI. Fibrinolysis

The clot is meant to be a temporary plug for hemostasis or wound healing, so there are natural mechanisms in the body for the efficient removal of fibrin. Various proteolytic enzymes and cells can dissolve fibrin depending on the circumstances, but the most specific mechanism involves the fibrinolytic system. The dissolution of fibrin clots under physiological conditions involves the binding of circulating plasminogen to fibrin, and the activation of plasminogen to the active protease, plasmin, by tissue-type plasminogen activator (tPA), also bound to fibrin (Fig. 1). Highly efficient, specific fibrinolysis requires the ternary complex, fibrin-plasmin(ogen)-t-PA; activation of plasminogen is further stimulated by the initial cleavage of fibrin (Thorsen, 1992; Wiman and Collen, 1978). Plasmin then cleaves fibrin at specific sites, yielding certain identifiable soluble fragments, as described above (Fig. 6) (Marder *et al.*, 1969).

Both plasminogen and t-PA bind to fibrin, but new binding sites are created through partial cleavage by plasmin to create new C-terminal lysine residues (Suenson *et al.*, 1984). This positive feedback mechanism accelerates lysis to ensure efficient degradation of the intact clot. As a result during lysis, by addition of t-PA to a clot, there is a region with a very high concentration of plasminogen and a very sharp lysis front and a narrow lysis zone, where degradation is occurring with movement of the fibers being degraded (Collet *et al.*, 2000; Sakharov and Rijken, 1995; Sakharov *et al.*, 1996). The lysine residues of partially digested fibrin bind to lysine binding sites in certain of the kringle domains of plasminogen and t-PA, but there are also lysine-independent binding sites, a finger domain in t-PA, and aminohexyl sites in plasminogen. Plasminogen binds near the end-to-end junction of two fibrin molecules, which accounts for the specificity of its binding to polymerizing fibrin rather than fibrinogen (Weisel *et al.*, 1994). γ-Chain residues, including γ312-324, bind t-PA; residues Aα148-160, with a lysine residue at 157, bind both t-PA and plasminogen with similar affinities (Nieuwenhuizen, 2001), but since the concentration of plasminogen is much higher in plasma, Aα148-160 serves as a plasminogen binding site. Structural studies revealed that Aα154-159

is in a solvent-inaccessible region of the coiled coil (Spraggon *et al.*, 1997), but it may be exposed during fibrin polymerization (Medved and Nieuwenhuizen, 2003; Yakovlev *et al.*, 2000). Similarly, there are sites that bind plasminogen and t-PA in the αC domain that are exposed in fibrin (Tsurupa and Medved, 2001). It has been suggested that t-PA initially binds to fibrin via its finger domain, while plasminogen binds via its aminohexyl sites, and then both molecules can bind via kringles as C-terminal lysines are generated (van Zonneveld *et al.*, 1986).

Antifibrinolytic compounds can block the conversion of plasminogen to plasmin, or directly bind to the active site of plasmin to inhibit fibrinolysis. The plasma protein, α2-macroglobulin, is a primary physiological inhibitor of plasmin. Plasmin released from fibrin is also very rapidly inactivated by α2-antiplasmin, which plays a role in the regulation of the fibrinolytic process (Aoki and Harpel, 1984). α2-antiplasmin inactivates plasmin in a very rapid reaction, interferes with plasminogen binding to fibrin, and is ligated to fibrin by Factor XIIIa (Sakata and Aoki, 1980). After α2-antiplasmin is covalently linked to fibrin's C-terminal α chain, it retains it ability to inhibit plasmin, a function that helps to prevent premature clot lysis.

Fibrinolysis is also inhibited by plasminogen activator inhibitor type 1, which is a 50 kDa serine protease inhibitor released from platelets and endothelial cells in response to various stimuli. Plasminogen activator inhibitor-1 binds to fibrin via high- or low-affinity sites and can inhibit t-PA also bound to the clot (Keijer *et al.*, 1991; Stringer and Pannekoek, 1995). This is an important pathway for limiting fibrinolysis, especially in platelet-rich thrombi and regions near endothelial cells.

Thrombin activable fibrinolysis inhibitor (TAFI) is a plasma protein that is activated by thrombin in the presence of thrombomodulin to a labile carboxypeptidase-B-like enzyme that inhibits fibrinolysis. When TAFIa is included in a clot undergoing lysis induced by tPA and plasminogen, the time to achieve lysis is prolonged and free lysine and arginine are released (Wang *et al.*, 1998). TAFIa retards the fibrin-enhanced activation of plasminogen by tPA and inhibits the accumulation of plasminogen at the lysis front (Sakharov *et al.*, 1997).

Fibrinolysis may also be regulated by lipoprotein(a), or Lp(a), which consists of low-density lipoprotein (a lipid core with apolipoprotein B-100 on the surface) and apolipoprotein(a) (Weisel *et al.*, 2001). Apolipoprotein(a) has great similarity to plasminogen, in that it has an (inactive) serine protease domain and a series of many kringles. Individuals have different concentrations of Lp(a) and isoforms with different numbers of kringles. Like plasminogen, Lp(a) binds tightly to fibrin, initially via the αC domains, and new binding sites are generated by partial digestion of

the clot (Tsurupa *et al.*, 2003). Lp(a) may compete with plasminogen, inhibiting its activation by t-PA and clot lysis (Angles-Cano *et al.*, 1994; Edelberg *et al.*, 1990; Harpel *et al.*, 1989; Loscalzo *et al.*, 1990). Other studies show that Lp(a) enhances binding of plasminogen and increases the rate of t-PA activation (Liu *et al.*, 1994).

VII. Fibrinogen Binding to Integrins in Platelet Aggregation and Other Cellular Interactions

Fibrin(ogen) binds specifically to integrin receptors on platelets, endothelial cells, and many other cells and plays a vital role in platelet aggregation and other aspects of cellular adhesion. Clot formation is a normal part of the wound healing process, sealing the injury and preventing bleeding. The fibrin clot serves as a scaffold for the migration of various cells, but less is known about this process than about earlier stages of healing. Fibroblasts migrate through the gel and deposit collagen, while macrophages and elements of the fibrinolytic system degrade the clot, with fibrin degradation products in turn acting on cells in the matrix.

Fibrinogen plays an important role in platelet aggregation (Bennett, 2001). The activity of circulating platelets is tightly controlled to prevent the spontaneous formation of platelet aggregates. Circulating platelets are inactive until they adhere to exposed subendothelial matrix proteins or are stimulated by soluble agonists such as ADP or thrombin. These activating events are associated with changes in platelet shape through reorganization of the platelet cytoskeleton, secretion of platelet granules, and an increase in the affinity of the integrin αIIbβ3 for soluble ligands, such as fibrinogen. Fibrinogen binding to platelets is responsible for platelet aggregation since it binds specifically to activated αIIbβ3, and the dimeric fibrinogen molecules bridge adjacent platelets (Fig. 9). The presence of platelets also has a dramatic effect on clot structure and mechanical properties (Fig. 7B) (Collet *et al.*, 2002).

Fibrinogen binds to stimulated platelets in a specific, calcium-dependent, saturable manner (Bennett, 2001). Residues in the γ chain of fibrinogen from 400–411 are necessary for binding to platelets. Fibrinogen also has two pairs of Arg-Gly-Asp sequences, in the Aα chain in the coiled-coil region (Aα 95–97) and in the carboxyl terminus (Aα 572–574). In spite of the fact that Arg-Gly-Asp sequences are a common motif for integrin binding, they do not appear to be important for platelet aggregation (Farrell *et al.*, 1992; Weisel *et al.*, 1992). Purified αIIbβ3 does not bind to the regions containing these sequences, but does bind to the distal ends of the fibrinogen (where the C-terminal γ chains are located). Furthermore, recombinant fibrinogen in which these sequences are mutated

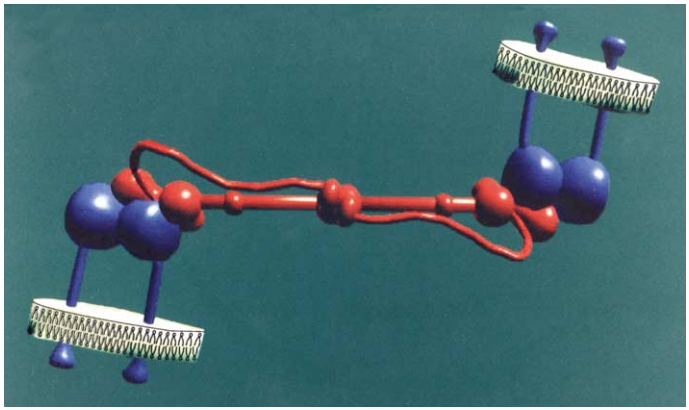


FIG. 9. Fibrinogen binding to the integrin αIIbβ3 in platelet aggregation. Diagram of fibrinogen-αIIbβ3 interactions, with fibrinogen in red and αIIbβ3 in blue. Two activated αIIbβ3 complexes, each with separated tails in the membrane of a different platelet, bind to the C-terminal γ chain at the ends of a fibrinogen molecule, such that the fibrinogen molecule forms a bridge between adjacent platelets. The shape and orientation of the proteins in the complex is based on structural studies. (Adapted from Weisel *et al.*, 1992.)

still supports platelet aggregation, whereas modification of the γ chain sequence eliminates platelet aggregation.

Studies with transgenic mice lacking the C-terminal Ala-Gly-Asp-Val sequence with recombinant human molecules lacking Ala-Gly-Asp–Val failed to support platelet aggregation but had normal clot retraction (Rooney *et al.*, 1996, 1998). These and other results suggest either that different receptors are responsible for platelet aggregation and clot retraction or the binding sites may be more complex (Remijn *et al.*, 2001).

Soluble fibrinogen also functions as a bridging molecule in other cell-cell and cell-extracellular matrix interactions in coagulation and inflammation. In addition, specific binding of fibrinogen to both integrin and nonintegrin proteins modulates the functions of many cells in a variety of receptor-mediated signal transduction processes. In addition to platelets, fibrinogen interacts with endothelial cells, leukocytes, fibroblasts, and epithelial cells.

Endothelial cell adhesion to surface-immobilized fibrinogen results in structural changes to the cells, including microfilament reorganization and cellular spreading. These interactions appear to be mediated by αvβ3 binding to the Arg-Gly-Asp sequence in the C-terminal Aα chain (Cheresh *et al.*, 1989; Dejana *et al.*, 1988; Francis *et al.*, 1993; Suehiro *et al.*, 1997). In

addition, fibrin missing both A and B fibrinopeptides promotes angiogenesis, which is important during wound healing and tumor growth. Generation of fibrin on top of an endothelial cell monolayer causes rapid rearrangement of endothelial cells into a capillary network (Martinez *et al.*, 2001). In this process, the N-terminal peptide β15-42 is involved in binding to VE cadherin on endothelial cells.

Fibrinogen also binds to the leukocyte integrin αMβ2/Mac-1, which is important for the inflammatory response and binds multiple ligands, and binding of fibrin(ogen) appears to be particularly important for leukocyte function. A variety of studies have identified the binding site for αMβ2/Mac-1 as being in the C-terminal γ-module, with γ377-395 and other sequences being involved (Lishko *et al.*, 2004; Ugarova *et al.*, 1998, 2003). To determine the significance of fibrin(ogen)-αMβ2 interaction *in vivo*, transgenic mice were generated in which the αMβ2 binding motif in the fibrinogen γ chain (390–396) was mutated, causing loss of αMβ2-mediated adhesion (Flick *et al.*, 2004). The result was a severely compromised inflammatory response, showing that integrin-fibrin(ogen) binding is critical to leukocyte function and that fibrinogen is important in regulating the inflammatory response.

Fibrinogen causes certain bacteria, such as *Staphylococcus aureus*, to clump. The C-terminal γ chain is necessary for the binding with the dimeric fibrinogen molecule forming bridges between the cells (Strong *et al.*, 1982). The clumping reactions are mediated by bacterial adhesins and are significant for bacterial colonization.

VIII. Dysfibrinogenemias, Variants, and Heterogeneity of Fibrinogen

Dysfibrinogenemias are characterized by structural changes in the fibrinogen molecule that result in detectable alterations in clotting or other properties of the molecule. Traditionally, they are named after the place of their discovery or where the patient lives. Most congenital defects are rare but offer the opportunity to study the effects of these molecular changes on fibrinogen function. These mutations are listed and described in several reviews (Galanakis, 1992, 1993; Matsuda and Sugo, 2001; McDonagh, 2000; Roberts *et al.*, 2001). In addition, a database of all mutations is maintained on the website http://www.geht.org/pages/database_fibro_uk.html.

Molecular defects that give rise to dysfibrinogenemias are commonly caused by single base mutations that lead to the substitution of a single amino acid. Other mutations can give rise to a stop codon, resulting in a

truncation of one of the chains. Finally, base deletions or additions may occur, with consequences for fibrinogen structure and function. Any such mutations can cause thrombosis, bleeding, or be asymptomatic, depending on their functional effects. They can affect fibrinopeptide cleavage, fibrin polymerization, crosslinking, fibrinolysis, or platelet aggregation. Since dysfibrinogenemias are discovered in a clinical setting, there are nearly always some functional effects.

In homozygous forms of the dysfibrinogenemias, the mutations occur in all fibrinogen molecules, so only abnormal homodimers will be present. However, homozygous mutations are very rare, so most dysfibrinogenemias are heterozygous. As a consequence, fibrinogen molecules in most subjects consist of various proportions of mutant homodimers, normal homodimers, and heterodimers.

Only a few examples of dysfibrinogenemias will be mentioned here. Some of the most common sites of the dysfibrinogens that have been identified are in the N-terminal Aα chain, in the regions involved in thrombin binding or cleavage or in the *A* knob. One frequently observed mutation is a substitution of Arg16 by His or Cys, with the former causing delayed release of fibrinopeptide A, but the latter completely blocking its release. If there is no fibrinopeptide A cleavage, fibrinopeptide B is released slowly and clots do form, but only at subambient temperatures (Shainoff and Dardik, 1979). Mutations of Gly17, Pro18, Arg19, or Val20 result in defective polymerization because the *A* site is altered. An Arg554 to Cys substitution in the Aα chain, with accompanying attachment of albumin to this free cysteine, leads to clots made up of very thin, slowly digested fibers, which probably accounts for the thrombosis that commonly is associated with this dysfibinogenemia (Collet *et al.*, 1993, 1996). Several Aα chain truncations have been described; clots from these mutants are generally made up of very thin fibers, again because of the lack of αC domains that normally enhance lateral aggregation. Several mutations, including AαArg141 to Ser, create a new consensus sequence for glycosylation, with the result that the additional mass and charge affect polymerization (Woodhead *et al.*, 1996).

Mutations of the Bβ chain are less common. Substitution of Cys for Gly at amino acid 15 results in prolonged polymerization from delayed release of fibrinopeptide B (Yoshida *et al.*, 1991). Mutation of Ala68 to Thr results in defective binding of thrombin to fibrin and consequent thrombosis (Koopman *et al.*, 1992). Deletion of residues 9–72, which includes a cysteine that is normally part of a disulfide bond, results in delayed release of both fibrinopeptides A and B (Liu *et al.*, 1985).

In the γ chain, many mutations in the C-terminal portion have been identified, particularly at position 275 (Cote *et al.*, 1998). This residue

makes contacts at the D:D interface, so that mutations affect fibrin polymerization. Several mutations have been identified in residues that are part of the calcium binding site in the γ chain. Mutations in or near the *a* pocket are common, but patients with most of these substitutions are asymptomatic, presumably because they are heterozygous. Finally, other mutations in the γ chain that affect polymerization may be in regions involved in lateral aggregation. In addition, some mutations in any chain may cause functional effects because of improper folding of a particular part of the polypeptide chain.

Some point mutations are responsible for congenital hypo- and afibrinogenemia, as a result of defects in molecular processing, assembly, secretion, and domain stability of fibrinogen. Afibrinogenemia is an autosomal recessive disorder characterized by the complete absence of detectable fibrinogen; analysis of the three fibrinogen genes in affected individuals has led to the identification of several causative mutations (Brennan *et al.*, 2001; Neerman-Arbez, 2001). Most of the cases identified so far were found to be caused by defective fibrinogen synthesis as a result of truncating mutations in the Aα chain gene. The graded severity of the hypo- and afibrinogenemias associated with homozygous Aα chain truncations suggest the minimal requirement for molecular assembly is the formation of the distal disulfide ring of the coiled coil. Heterozygous mutations of certain residues that perturb the five-stranded beta sheet of the D region are absent from plasma fibrinogen. Other mutations can affect intracellular proteolysis, chain assembly, and export.

As an animal model of afibrinogenemia, to examine the role of fibrinogen in hemostasis, wound repair, development, and disease pathogenesis, the fibrinogen Aα chain gene was disrupted in mice (Suh *et al.*, 1995). Homozygous, Aα chain-deficient (A$\alpha-/-$) mice were born normal in appearance, and there was no evidence of fetal loss of these animals. None of the chains of fibrinogen were immunologically detectable in the circulation of adult A$\alpha-/-$ mice, and blood samples did not clot or support platelet aggregation *in vitro*. Although overt bleeding events developed shortly after birth in about 30% of A$\alpha-/-$ mice, most newborns eventually controlled the loss of blood. Juveniles and young adult A$\alpha-/-$ mice were predisposed to spontaneous, fatal abdominal hemorrhage, but long-term survival was variable. Pregnancy uniformly resulted in fatal uterine bleeding around the tenth day of gestation. These results are consistent with human afibrinogenemic or severe hypofibrinogenemic patients, in whom gestation invariably fails during the first trimester. In these patients, gestation is salvaged by fibrinogen replacement throughout pregnancy, indicating the absolute need for fibrinogen and a dynamic fibrin deposition and removal process for normal placentation (Galanakis *et al.*, 1996).

Additionally, it was found that fibrin(ogen) is important for wound healing in cellular migration, organization, and establishment of wound stability and strength (Drew *et al.*, 2001). Spontaneous bloodborne and lymphatic metastasis, but not primary tumor growth or angiogenesis, was diminished in fibrinogen-deficient mice (Palumbo *et al.*, 2002). Plasminogen deficiency in mice results in severe thrombosis, in addition to delayed wound healing, spontaneous gastrointestinal ulceration, rectal prolapse, and wasting (Bugge *et al.*, 1995; Romer *et al.*, 1996). Mice deficient in both plasminogen and fibrinogen demonstrate that removal of fibrin(ogen) from the extracellular environment alleviates the diverse spontaneous pathologies associated with plasminogen deficiency and corrects healing times (Bugge *et al.*, 1996).

There are also some acquired dysfibrinogenemias from alterations to fibrinogen, usually as a result of disease processes. Liver disease (including cirrhosis, hepatitis, or hepatic carcinomas) is the most common cause of these modifications. In these cases, there is usually an increase in the sialic acid content of the fibrinogen's carbohydrate (Martinez *et al.*, 1983).

Variants of fibrinogen are also present as a result of several common polymorphisms. The most widespread polymorphisms in the fibrinogen gene occur in the noncoding regions and can result in changes in plasma fibrinogen levels. Two common polymorphisms cause a change in the amino acid sequence of fibrinogen (Baumann and Henschen, 1993). In the Aα chain, a change from 312 Thr to Ala occurs at a frequency of about 23%. The other common polymorphism occurs at position 448 in the Bβ chain with substitution of Arg to Lys, at a frequency of 15%. Associations between disease and each of these polymorphisms have been established (Ariëns *et al.*, 2002). The Aα312 polymorphism affects clot structure and properties (Standeven *et al.*, 2003).

There are many molecular forms of fibrinogen present in blood, as seen from variations in biochemical properties and gel electrophoretic behavior. It has been calculated that fibrinogen may occur in more than a million nonidentical forms in a healthy individual as a result of the many combinations of modified or polymorphic sites (Henschen-Edman, 2001). Several of these variations—such as sequence polymorphisms, carbohydrate content and structure, and splice variants—have already been mentioned. In addition, there is variation in phosphorylation at a few specific serine sites, proline hydroxylation, tyrosine sulfation, asparagine or glutamine deamidation, glutamine cyclization, and methionine oxidation. The C-terminal Aα chain is particularly susceptible to digestion by proteolytic enzymes, but some digestion also commonly occurs at specific sites in the β and γ chains, and consequently lower molecular weight forms are nearly always present in plasma fibrinogen. Abnormal variants

occur in patients with several conditions, including glycation of lysine residues in diabetes mellitus, crosslinked degradation products in disseminated intravascular coagulation and some types of cancer, and antibody complexes in certain autoimmune diseases. Tobacco smokers tend to have nitration of tyrosine and oxidation of methionine, histidine, or tryptophane residues, and individuals treated with acetylsalicylic acid have acetylated lysine residues. All of these polymorphic forms may show considerable differences in their functional properties.

IX. Evolution of Fibrinogen and Fibrinogen-Like Domains

The thrombin-catalyzed clotting of fibrinogen to form a fibrin gel is common to all extant vertebrates. Because clots are necessary to stop bleeding, but can also cause thrombosis if not dissolved in a timely fashion, an effective scheme for fibrinolysis evolved concomitantly. The amino acid sequences of the Aα, Bβ, and γ chains are homologous, indicating that they have evolved from a common ancestor (Doolittle *et al.*, 1997; Henschen *et al.*, 1982). This homology even extends to the genomic sequences, with two intron/exon boundaries being conserved in all three chains; another one is common to both Bβ and γ chains (Medved, 1990). Taken together, all of this evidence suggests that the evolution of the three chains from a common ancestor through a series of duplications and inversions began about a billion years ago. It has been suggested that an ancestral gene duplicated to produce the α chain gene and a pre $\beta-\gamma$ gene (Doolittle *et al.*, 1997; Henschen *et al.*, 1982). The pre $\beta-\gamma$ gene then duplicated about 500 million years ago to form the β and γ genes. Zebrafish have all major hemostatic proteins, allowing some evolutionary predictions (Hanumanthaiah *et al.*, 2002) and thrombin has been used as a probe to investigate fibrinogen to fibrin conversion at different stages of embryonic development (Jagadeeswaran and Liu, 1997).

Amino acid sequence comparisons imply that fibrinogen evolved before the divergence of vertebrates and invertebrates, so it is reasonable to suppose that a protein with domains similar to those of fibrinogen might exist in invertebrates. A fibrinogen-like sequence was identified in cDNA prepared from the soft tissues of a sea cucumber, *Parastichopus parvimensis* (Xu and Doolittle, 1990). Two proteins corresponding to the cloned mRNAs, FReP-A and FReP-B, have sequences with similarity to the C-terminal two-thirds of vertebrate fibrinogen Bβ and γ chains. Comparisons of various fibrinogen-related sequences suggest that the sea cucumber proteins diverged before the β-γ gene duplication.

Nucleotide sequences of fibrinogen-like proteins have been identified in a number of other genes, including the tenascin family of extracellular matrix proteins (Erickson, 1994; Jones *et al.*, 1988), fibroleukin (Koyama *et al.*, 1987), and a protein encoding *scabrous* of the developing *Drosophila* eye (Baker *et al.*, 1990; Lee *et al.*, 1998). The existence of these fibrinogen-like domains suggests that the specific and tightly-controlled intermolecular interactions that occur in fibrin polymerization and in platelet aggregation may be used in other aspects of cellular function and developmental biology.

X. Conclusions

There is now much structural information about fibrinogen, including its primary sequence, the chain connectivity via disulfide bonds, and organization into domains. There are X-ray crystallographic structures of large parts of the molecule, but some key regions have not yet been visualized. Some aspects of calcium binding are beginning to be discovered, but less is known of the low affinity sites.

Tremendous strides have been taken in understanding the metabolism and biosynthesis of fibrinogen, but there is still much to be done. Almost nothing is known of the catabolism of fibrinogen. The functions of the alternately spliced forms and the many posttranslational modifications are only partly known. Many dysfibrinogenemias have been discovered, but most have only been relatively superficially characterized.

This is an exciting time for studies of fibrin polymerization. The basics have been outlined and it seems that we may be tantalizingly close to achieving a real understanding of the molecular mechanisms. Of the complementary binding sites, the holes have been seen but not the knobs. The holes with specific peptides have also been visualized, but not the protein-protein interactions in fibrin. Little is known of either the *B* or *b* interactions, or of the function of the βC domain. Structural changes during polymerization are only beginning to be discovered. Binding interactions in lateral aggregation are unknown.

The mechanical properties of many types of clots have been measured, but the origin of clot viscoelasticity is a mystery. Factor XIIIa-induced crosslinking sites have been identified in the primary sequence, but the structure of the crosslinked C-terminal γ chains is a matter of debate and that of the α chains is unknown. The biochemistry of fibrinolysis has been determined, but less is known of the physical mechanisms involved or of the interactions of regulatory systems.

Binding interactions of several plasma proteins to fibrin(ogen) have been identified, but the functional consequences of most of them are not

well characterized. The interactions of fibrinogen with integrins on cells is an extremely active field of study that was only touched upon briefly in this review. Fibrinogen binding to platelets continues to be a model system for studies of cell adhesion.

The involvement of fibrinogen in biological functions beyond hemostasis and thrombosis is a wide open field. Nearly all research has been carried out with human fibrinogen, with isolated studies of a few other mammals, frog, chicken, zebrafish, and lamprey. There have been few comparative and evolutionary studies.

Acknowledgments

This work was supported by NIH grant HL30954. The author thanks Rustem Litvinov, Dennis Galanakis, Leonid Medved, Robert Ariëns, Oleg Gorkun, and Leona Mášová for providing many helpful suggestions and comments on this review, and Igor Pechik for generating Fig. 4.

References

Amrani, D. L., Plant, P. W., Pindyck, J., Mosesson, M. W., and Grieninger, G. (1983). Structural analysis of fibrinogen synthesized by cultured chicken hepatocytes in the presence or absence of dexamethasone. *Biochim. Biophys. Acta.* **743,** 394–400.

Angles-Cano, E., Hervio, L., Rouy, D., Fournier, C., Chapman, J. M., Laplaud, M., and Koschinsky, M. L. (1994). Effects of lipoprotein(a) on the binding of plasminogen to fibrin and its activation by fibrin-bound tissue-type plasminogen activator. *Chem. Phys. Lipids* **67–68,** 369–380.

Aoki, N., and Harpel, P. C. (1984). Inhibitors of the fibrinolytic enzyme system. *Semin. Thromb. Hemost.* **10,** 24–41.

Ariëns, R. A. S., Lai, T.-S., Weisel, J. W., Greenberg, C. S., and Grant, P. J. (2002). Role of Factor XIII in fibrin clot formation and effects of genetic polymorphisms. *Blood* **100,** 743–754.

Bacon-Baguley, T., Ogilvie, M. L., Gartner, T. K., and Walz, D. A. (1990). Thrombospondin binding to specific sequences within the A alpha- and B beta-chains of fibrinogen. *J. Biol. Chem.* **265,** 2317–2323.

Bailey, K., Astbury, W. T., and Rudall, K. M. (1943). Fibrinogen and fibrin as members of the keratin-myosin group. *Nature* **151,** 716–717.

Baker, N. E., Mlodzik, M., and Rubin, G. M. (1990). Spacing differentiation in the developing *Drosophila* eye: A fibrinogen-related lateral inhibitor encoded by scabrous. *Science* **250,** 1370–1377.

Baradet, T. C., Haselgrove, J. C., and Weisel, J. W. (1995). Three-dimensional reconstruction of fibrin clot networks from stereoscopic intermediate voltage electron microscope images and analysis of branching. *Biophys. J.* **68,** 1551–1560.

Baumann, R. E., and Henschen, A. H. (1993). Human fibrinogen polymorphic site analysis by restriction endonuclease digestion and allele-specific polymerase chain reaction amplification: Identification of polymorphisms at positions A alpha 312 and B beta 448. *Blood* **82,** 2117–2124.

Baumhueter, S., Courtois, G., Morgan, J. G., and Crabtree, G. R. (1989). The role of HNF-1 in liver-specific gene expression. *Ann. NY Acad. Sci.* **557,** 272–279.

Beguin, S., and Kumar, R. (1997). Thrombin, fibrin and platelets: A resonance loop in which von Willebrand factor is a necessary link. *Thromb. Haemost.* **78,** 590–594.

Bennett, J. S. (2001). Platelet-fibrinogen interactions. *Ann. NY Acad. Sci.* **936,** 340–354.

Bernocco, S., Ferri, F., Profumo, A., Cuniberti, C., and Rocco, M. (2000). Polymerization of rod-like macromolecular monomers studied by stopped-flow, multiangle light scattering: Set-up, data processing, and application to fibrin formation. *Biophys. J.* **79,** 561–583.

Blombäck, B. (1991). Fibrinogen and Fibrin Formation and Its Role in Fibrinolysis. *In* ''Biotechnology of Blood'' (J. Goldstein, Ed.), pp. 225–279. Butterworth-Heinemann, Boston.

Blombäck, B. (2001). Fibrinogen: Evolution of the structure-function concept. *Ann. NY Acad. Sci.* **936,** 1–11.

Blombäck, B., and Blombäck, M. (1956). Purification of human and bovine fibrinogen. *Arkiv. Kemi.* **10,** 415–443.

Blombäck, B., Hessel, B., Hogg, D., and Therkildsen, L. (1978). A two-step fibrinogen-fibrin transition in blood coagulation. *Nature* **275,** 501–505.

Blombäck, B., Carlsson, K., Fatah, K., Hessel, B., and Procyk, R. (1994). Fibrin in human plasma: Gel architectures governed by rate and nature of fibrinogen activation. *Thromb. Res.* **75,** 521–538.

Bode, W., and Stubbs, M. T. (1993). Spatial structure of thrombin as a guide to its multiple sites of interaction. *Semin. Thromb. Hemost.* **19,** 321–333.

Brennan, S. O., Fellowes, A. P., and George, P. M. (2001). Molecular mechanisms of hypo- and afibrinogenemia. *Ann. NY Acad. Sci.* **936,** 91–100.

Brown, J. H., Volkmann, N., Jun, G., Henschen, A. H., and Cohen, C. (2000). The crystal structure of modified bovine fibrinogen. *Proc. Natl. Acad. Sci. USA* **97,** 85–90.

Budzynski, A. Z. (1986). Fibrinogen and fibrin: Biochemistry and pathophysiology. *Crit. Rev. Oncol.-Hematol.* **6,** 97–146.

Bugge, T. H., Flick, M. J., Daugherty, C. C., and Degen, J. L. (1995). Plasminogen deficiency causes severe thrombosis but is compatible with development and reproduction. *Genes Develop.* **9,** 794–807.

Bugge, T. H., Kombrinck, K. W., Flick, M. J., Daugherty, C. C., Danton, M. J., and Degen, J. L. (1996). Loss of fibrinogen rescues mice from the pleiotropic effects of plasminogen deficiency. *Cell* **87,** 709–719.

Carr, M. E. (1988). Fibrin formed in plasma is composed of fibers more massive than those formed from purified fibrinogen. *Thromb. Haemost.* **59,** 535–539.

Carr, M. E., and Hardin, C. L. (1987). Fibrin has larger pores when formed in the presence of erythrocytes. *Am. J. Physiol.* **253,** H1069–H1073.

Carr, M. E., Gabriel, D. A., and McDonagh, J. (1986). Influence of Ca2+ on the structure of reptilase-derived and thrombin-derived fibrin gels. *Biochem. J.* **239,** 513–516.

Carr, M. E., Kaminski, M., McDonagh, J., and Gabriel, D. A. (1985). Influence of ionic strength, peptide release, and calcium ion on the structure of reptilase and thrombin-derived gels. *Thromb. Haemost.* **54,** 159–165.

Cheresh, D. A., Berliner, S. A., Vicente, V., and Ruggeri, Z. M. (1989). Recognition of distinct adhesive sites on fibrinogen by related integrins on platelets and endothelial cells. *Cell* **58,** 945–953.

Chung, D. W., and Davie, E. W. (1984). Gamma and gamma′ chains of human fibrinogen are produced by alternative RNA splicing. *Biochem.* **23,** 4232–4236.

Chung, D. W., Harris, J. E., and Davie, E. W. (1990). Nucleotide sequences of the three genes coding for human fibrinogen. *Adv. Exp. Med. Biol.* **281,** 39–48.

Chung, D. W., Rixon, M. W., Que, B. G., and Davie, E. W. (1983). Cloning of fibrinogen genes and their cDNA. *Ann. NY Acad. Sci.* **408,** 449–456.

Cohen, C., Revel, J.-P., and Kucera, J. (1963). Paracrystalline forms of fibrinogen. *Science* **141,** 436–438.

Cohen, C., Weisel, J. W., Phillips, G. N., Stauffacher, C. V., Fillers, J. P., and Daub, E. (1983). The structure of fibrinogen and fibrin: I. Electron microscopy and X-ray crystallography of fibrinogen. *Ann. NY Acad. Sci.* **408,** 194–213.

Cohn, E. J., Strong, L. E., Hughes, W. L., Mulford, D. J., Ashworth, J. N., Melin, M., and Taylor, H. L. (1946). Preparation and properties of serum and plasma proteins. IV. A system for the separation into fractions of the protein and lipoprotein components of biological tissues and fluids. *J. Am. Chem. Soc.* **68,** 459–475.

Collet, J.-P., Montalescot, G., Lesty, C., and Weisel, J. W. (2002). A structural and dynamic investigation of the facilitating effect of glycoprotein IIb/IIIa inhibitors in dissolving platelet-rich clots. *Circ. Res.* **90,** 428–434.

Collet, J.-P., Woodhead, J. L., Soria, J., Soria, C., Mirshahi, M. C., Caen, J. P., and Weisel, J. W. (1996). Fibrinogen Dusart: Electron microscopy of molecules, fibers, and clots and viscoelastic measurements. *Biophys. J.* **70,** 500–510.

Collet, J.-P., Park, D., Lesty, C., Soria, J., Soria, C., Montalescot, G., and Weisel, J. W. (2000). Influence of fibrin network conformation and fibrin fiber diameter on fibrinolysis speed: Dynamic and structural approaches by confocal microscopy. *Arterioscler. Thromb. Vasc. Biol.* **20,** 1354–1361.

Collet, J. P., Soria, J., Mirshahi, M., Hirsch, M., Dagonnet, F. B., Caen, J., and Soria, C. (1993). Dusart syndrome: A new concept of the relationship between fibrin clot architecture and fibrin clot degradability: Hypofibrinolysis related to an abnormal clot structure. *Blood* **82,** 2462–2469.

Cote, H. C., Lord, S. T., and Pratt, K. P. (1998). Gamma-chain dysfibrinogenemias: Molecular structure-function relationships of naturally occurring mutations in the gamma chain of human fibrinogen. *Blood* **92,** 2195–2212.

Crabtree, G. R. (1987). The molecular biology of fibrinogen. *In* ''The Molecular Basis of Blood Diseases'' (G. Stamatoyannopoulos, A. W. Nienhuis, P. Leder, and P. W. Majerus, Eds.), pp. 631. W. B. Saunders, Philadelphia.

Dang, C. V., Shin, C. K., Bell, W. R., Nagaswami, C., and Weisel, J. W. (1989). Fibrinogen sialic acid residues are low affinity calcium-binding sites that influence fibrin assembly. *J. Biol. Chem.* **264,** 15104–15108.

De Cristofaro, R., and Di Cera, E. (1992). Modulation of thrombin-fibrinogen interaction by specific ion effects. *Biochem.* **31,** 257–265.

Dejana, E., Languino, L. R., Colella, S., Corbascio, G. C., Plow, E. F., Ginsberg, M. H., and Marchisio, P. C. (1988). The localization of a platelet GpIIb-IIIa-related protein in endothelial cell adhesion structures. *Blood* **71,** 566–572.

Di Stasio, E., Nagaswami, C., Weisel, J. W., and Di Cera, E. (1998). Cl-regulates the structure of the fibrin clot. *Biophys. J.* **75,** 1973–1979.

Dixit, V. M., Grant, G. A., Frazier, W. A., and Santoro, S. A. (1984). Isolation of the fibrinogen-binding region of platelet thrombospondin. *Biochem. Biophys. Res. Commun.* **119,** 1075–1081.

Donovan, J. W., and Mihalyi, E. (1985). Clotting of fibrinogen. 1. Scanning calorimetric study of the effect of calcium. *Biochem.* **24,** 3434–3443.

Doolittle, R. F. (1973). Structural aspects of the fibrinogen to fibrin conversion. *Adv. Protein Chem.* **27,** 1–109.

Doolittle, R. F. (1984). Fibrinogen and fibrin. *Annu. Rev. Biochem.* **53,** 195–229.

Doolittle, R. F., Goldbaum, D. M., and Doolittle, L. R. (1978). Designation of sequences involved in the ''coiled-coil'' interdomainal connections in fibrinogen: Constructions of an atomic scale model *J. Mol. Biol.* **120,** 311–325.

Doolittle, R. F., Spraggon, G., and Everse, S. J. (1997). Evolution of vertebrate fibrin formation and the process of its dissolution. *Ciba. Found. Symp.* **212,** 4–23.

Drew, A. F., Liu, H., Davidson, J. M., Daugherty, C. C., and Degen, J. L. (2001). Wound-healing defects in mice lacking fibrinogen. *Blood* **97,** 3691–3698.

Dyr, J. E., Blombäck, B., Hessel, B., and Kornalik, F. (1989). Conversion of fibrinogen to fibrin induced by preferential release of fibrinopeptide B. *Biochim Biophys. Acta.* **990,** 18–24.

Edelberg, J. M., Gonzalez-Gronow, M., and Pizzo, S. V. (1990). Lipoprotein(a) inhibition of plasminogen activation by tissue-type plasminogen activator. *Thromb. Res.* **57,** 155–162.

Erickson, H. P. (1994). Evolution of the tenascin family–Implications for function of the C-terminal fibrinogen-like domain. *Perspectives Develop. Neurobiol.* **2,** 9–19.

Erickson, H. P., and Fowler, W. E. (1983). Electron microscopy of fibrinogen, its plasmic fragments and small polymers. *Ann. NY Acad. Sci.* **408,** 146–163.

Everse, S. J., Spraggon, G., Veerapandian, L., and Doolittle, R. F. (1999). Conformational changes in fragments D and double-D from human fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide. *Biochem.* **38,** 2941–2946.

Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M., and Doolittle, R. F. (1998). Crystal structure of fragment double-D from human fibrin with two different bound ligands. *Biochem.* **37,** 8637–8642.

Farrell, D. H., Thiagarajan, P., Chung, D. W., and Davie, E. W. (1992). Role of fibrinogen alpha and gamma chain sites in platelet aggregation. *Proc. Natl. Acad. Sci. USA* **89,** 10729–10732.

Fatah, K., Silveira, A., Tornvall, P., Karpe, F., Blombäck, M., and Hamsten, A. (1996). Proneness to formation of tight and rigid fibrin gel structures in men with myocardial infarction at a young age. *Thromb. Haemost.* **76,** 535–540.

Ferri, F., Greco, M., Arcovito, G., De Spirito, M., and Rocco, M. (2002). Structure of fibrin gels studied by elastic light scattering techniques: Dependence of fractal dimension, gel crossover length, fiber diameter, and fiber density on monomer concentration. *Phys. Rev. E.* **66**(1 Pt 1), 011913.

Ferri, F., Greco, M., Arcovito, G., Bassi, F. A., De Spirito, M., Paganini, E., and Rocco, M. (2001). Growth kinetics and structure of fibrin gels. *Phys. Rev. E.* **63**(3 Pt 1), 031401.

Ferry, J. D. (1988). Structure and rheology of fibrin networks. *In* ''Biological and Synthetic Polymer Networks'' (O. Kramer, Ed.), pp. 41–55. Elsevier, Amsterdam.

Flick, M. J., Du, X., Witte, D. P., Jirouskova, M., Soloviev, D. A., Busuttil, S. J., Plow, E. F., and Degen, J. L. (2004). Leukocyte engagement of fibrin(ogen) via the integrin receptor alphaMbeta2/Mac-1 is critical for host inflammatory response *in vivo. J. Clin. Invest.* **113,** 1596–1606.

Fowler, W. E., Hantgan, R. R., Hermans, J., and Erickson, H. P. (1981). Structure of the fibrin protofibril. *Proc. Natl. Acad. Sci. USA* **78,** 4872–4876.

Francis, C. W., Bunce, L. A., and Sporn, L. A. (1993). Endothelial cell responses to fibrin mediated by FPB cleavage and the amino terminus of the beta chain. *Blood Cells* **19,** 306–307.

Freyssinet, J. M., Torbet, J., Hudry-Clergeon, G., and Maret, G. (1983). Fibrinogen and fibrin structure and fibrin formation measured by using magnetic orientation. *Proc. Natl. Acad. Sci. USA* **80,** 1616–1620.

Fu, Y., and Grieninger, G. (1994). Fib420: A normal human variant of fibrinogen with two extended alpha chains. *Proc. Natl. Acad. Sci. USA* **91,** 2625–2628.

Galanakis, D. K. (1992). Fibrinogen anomalies and disease: A clincal update. *Hematol. Oncol. Clin. N. Am.* **6,** 1171–1187.

Galanakis, D. K. (1993). Inherited dysfibrinogenemia: Emerging abnormal structure associations with pathologic and nonpathologic dysfunctions. *Semin. Thromb. Hemost.* **19,** 386–395.

Galanakis, D. K., Lane, B. P., and Simon, S. R. (1987). Albumin modulates lateral assembly of fibrin polymers: Evidence of enhanced fine fibril formation and of unique synergism with fibrinogen. *Biochem.* **26,** 2389–2400.

Galanakis, D. K., Nuovo, G., Spitzer, S., Kaplan, C., and Scharrer, I. (1996). Fibrinogen mRNA and antigen co-present in human trophoblasts *in situ*: Possible implications. *Thromb. Res.* **81,** 263–269.

Godyna, S., Diaz-Ricart, M., and Argraves, W. S. (1996). Fibulin-1 mediates platelet adhesion via a bridge of fibrinogen. *Blood* **88,** 2569–2577.

Gorkun, O. V., Veklich, Y. I., Medved', L. V., Henschen, A. H., and Weisel, J. W. (1994). Role of the αC domains in fibrin formation. *Biochemistry* **33,** 6986–6997.

Greenberg, C. S., Miraglia, C. C., Rickles, F. R., and Shuman, M. A. (1985). Cleavage of blood coagulation factor XIII and fibrinogen by thrombin during *in vitro* clotting. *J. Clin. Invest.* **75,** 1463–1470.

Greenberg, C. S., Lai, T.-S., Ariens, R. A. S., Weisel, J. W., and Grant, P. J. (2003). Biology and Disorders of Fibrinogen and Factor XIII. *In* ''Blood. Principles and Practice of Hematology'' (R. I. Handin, S. E. Lux, and T. P. Stossel, Eds.), pp. 1225–1248. Lippincott Williams & Wilkins, Philadelphia.

Guinto, E. R., Vindigni, A., Ayala, Y. M., Dang, Q. D., and Di Cera, E. (1995). Identification of residues linked to the slow–>fast transition of thrombin. *Proc. Natl. Acad. Sci. USA* **92,** 11185–11189.

Hada, M., Kaminski, M., Bockenstedt, P., and McDonagh, J. (1986). Covalent cross-linking of von Willebrand factor to fibrin. *Blood* **68,** 95–101.

Haidaris, P. J. (1997). Induction of fibrinogen biosynthesis and secretion from cultured pulmonary epithelial cells. *Blood* **89,** 873–882.

Haidaris, P. J., and Courtney, M. A. (1990). Tissue-specific and ubiquitous expression of fibrinogen gamma-chain mRNA. *Blood Coagul. Fibrinolysis* **1,** 433–437.

Haidaris, P. J., Francis, C. W., Sporn, L. A., Arvan, D. S., Collichio, F. A., and Marder, V. J. (1989). Megakaryocyte and hepatocyte origins of human fibrinogen biosynthesis exhibit hepatocyte-specific expression of gamma chain-variant polypeptides. *Blood* **74,** 743–750.

Hall, C. E., and Slayter, H. S. (1959). The fibrinogen molecule: Its size, shape, and mode of polymerization. *J. Biophys. Biochem. Cytol.* **5,** 11–16.

Hanna, L. S., Scheraga, H. A., Francis, C. W., and Marder, V. J. (1984). Comparison of structures of various human fibrinogens and a derivative thereof by a study of the kinetics of release of fibrinopeptides. *Biochem.* **23,** 4681–4687.

Hantgan, R. R., and Hermans, J. (1979). Assembly of fibrin. A light scattering study. *J. Biol. Chem.* **254,** 11272–11281.

Hantgan, R. R., Fowler, W. E., Erickson, H. P., and Hermans, J. (1980). Fibrin assembly: A comparison of electron microscopic and light scattering results. *Thromb. Haemost.* **44,** 119–124.

Hantgan, R. R., Simpson-Haidaris, P. J., Francis, C. W., and Marder, V. J. (2000). Fibrinogen structure and physiology. *In* ''Hemostasis and Thrombosis: Basic Principles and Clinical Practice'' (R. W. Colman, J. Hirsh, V. J. Marder, A. W. Clowes, and J. N. George, Eds.), pp. 203–232. Lippincott, Williams & Wilkins, Philadelphia.

Hanumanthaiah, R., Day, K., and Jagadeeswaran, P. (2002). Comprehensive analysis of blood coagulation pathways in teleostei: Evolution of coagulation factor genes and identification of zebrafish factor VIIi. *Blood Cells* **29,** 57–68.

Harpel, P. C., Gordon, B. R., and Parker, T. S. (1989). Plasmin catalyzes binding of lipoprotein(a) to immobilized fibrinogen and fibrin. *Proc. Natl. Acad. Sci. USA* **86,** 3847–3851.

Harrison, P., Wilbourn, B., Debili, N., Vainchenker, W., Breton-Gorius, J., Lawrie, A. S., Masse, J. M., Savidge, G. F., and Cramer, E. M. (1989). Uptake of plasma fibrinogen into the alpha granules of human megakaryocytes and platelets. *J. Clin. Invest.* **73,** 1123–1129.

Henschen, A., and McDonagh, J. (1986). Fibrinogen, fibrin and factor XIII. *In* ''Blood Coagulation'' (R. F. A. Zwaal, and H. C. Hemker, Eds.), pp. 171–234. Elsevier Science, Amsterdam.

Henschen, A. H., Lottspeich, F., Kehl, M., Southan, C., and Lucas, J. (1982). Structure-function-evolution relationship in fibrinogen. *In* ''Fibrinogen. Recent Biochemical and Medical Aspects'' (A. H. Henschen, H. Graeff, and F. Lottspeich, Eds.), pp. 67–82. Walter de Gruyter, Berlin.

Henschen-Edman, A. H. (2001). Fibrinogen non-inherited heterogeneity and its relationship to function in health and disease. *Ann. NY Acad. Sci.* **936,** 580–593.

Hirose, S., Oda, K., and Ikehara, Y. (1988). Biosynthesis, assembly, and secretion of fibrinogen in cultured rat hepatocytes. *Biochem. J.* **251,** 373–377.

Hogan, K. A., Bolliger, B., Okumura, N., and Lord, S. T. (2001). The formation of beta fibrin requires a functional a site. *Ann. NY Acad. Sci.* **936,** 331–334.

Hu, C. H., Harris, J. E., Davie, E. W., and Chung, D. W. (1995). Characterization of the 5′-flanking region of the gene for the alpha chain of human fibrinogen. *J. Biol. Chem.* **270,** 28342–28349.

Jagadeeswaran, P., and Liu, Y. C. (1997). Developmental expression of thrombin in zebrafish embryos: A novel model to study hemostasis. *Blood Cells* **23,** 147–156.

Jandrot-Perrus, M., Mosesson, M. W., M.H., D., and Menache, D. (1979). Studies of platelet fibrinogen from a subject with a congenital plasma fibrinogen abnormality (fibrinogen Paris I). *Blood* **54,** 1109–1116.

Janmey, P. A., Amis, E. J., and Ferry, J. D. (1983). Rheology of fibrin clots. VI. Stress relaxation, creep, and differential dynamic modulus of fine clots in large shearing deformations. *J. Rheol.* **27,** 135–153.

Jones, F. S., Burgoon, M. P., Hoffman, S., Crossin, K. L., Cunningham, B. A., and Edelman, G. M. (1988). A cDNA clone for cytotactin contains sequences similar to epidermal growth factor-like repeats and segments of fibronectin and fibrinogen. *Proc. Natl. Acad. Sci. USA* **85,** 2186–2190.

Kant, J. A., Fornace, A. J., Saxe, D., Simon, M. I., McBride, O. W., and Crabtree, G. R. (1985). Evolution and organization of the fibrinogen locus on chromosome 4: Gene duplication accompanied by transposition and inversion. *Proc. Natl. Acad. Sci. USA* **185,** 1–19.

Keijer, J., Linders, M., van Zonneveld, A. J., Ehrlich, H. J., de Boer, J. P., and Pannekoek, H. (1991). The interaction of plasminogen activator inhibitor 1 with plasminogen activators (tissue-type and urokinase-type) and fibrin: Localization of interaction sites and physiologic relevance. *Blood* **78,** 401–409.

Koopman, J., Haverkate, F., Lord, S. T., Grimbergen, J., and Mannucci, P. M. (1992). Molecular basis of fibrinogen Naples associated with defective thrombin binding and thrombophilia. Homozygous substitution of B beta 68 Ala —> Thr. *J. Clin. Invest.* **90,** 238–244.

Kostelansky, M. S., Betts, L., Gorkun, O. V., and Lord, S. T. (2002). 2.8 A crystal structures of recombinant fibrinogen fragment D with and without two peptide ligands: GHRP binding to the ''b'' site disrupts its nearby calcium-binding site. *Biochem.* **41,** 12124–12132.

Kostelansky, M. S., Bolliger-Stucki, B., Betts, L., Gorkun, O. V., and Lord, S. T. (2004a). B beta Glu397 and B beta Asp398 but not B beta Asp432 are required for ''B:b'' interactions. *Biochem.* **43,** 2465–2474.

Kostelansky, M. S., Lounes, K. C., Ping, L. F., Dickerson, S. K., Gorkun, O. V., and Lord, S. T. (2004b). Calcium-binding site beta 2, adjacent to the ''b'' polymerization site, modulates lateral aggregation of protofibrils during fibrin polymerization. *Biochem.* **43,** 2475–2483.

Koyama, T., Hall, L. R., Haser, W. G., Tonegawa, S., and Saito, H. (1987). Structure of a cytotoxic T-lymphocyte-specific gene shows a strong homology to fibrinogen beta and gamma chains. *Proc. Natl. Acad. Sci. USA* **84,** 1609–1613.

Laki, K. (1951). The polymerization of proteins: The action of thrombin on fibrinogen. *Arch. Biochem. Biophys* **32,** 317–324.

Langer, B. G., Weisel, J. W., Dinauer, P. A., Nagaswami, C., and Bell, W. R. (1988). Deglycosylation of fibrinogen accelerates polymerization and increases lateral aggregation of fibrin fibers. *J. Biol. Chem.* **263,** 15056–15063.

Lee, E. C., Yu, S. Y., Hu, X., Mlodzik, M., and Baker, N. E. (1998). Functional analysis of the fibrinogen-related scabrous gene from *Drosophila melanogaster* identifies potential effector and stimulatory protein domains. *Genetics* **150,** 663–673.

Leung, L. L. (1984). Role of thrombospondin in platelet aggregation. *J. Clin. Invest.* **74,** 1764–1772.

Lewis, S. D., Shields, P. P., and Shafer, J. A. (1985). Characterization of the kinetic pathway for liberation of fibrinopeptides during assembly of fibrin. *J. Biol. Chem.* **260,** 10192–10199.

Lishko, V. K., Podolnikova, N. P., Yakubenko, V. P., Yakovlev, S., Medved, L., Yadav, S. P., and Ugarova, T. P. (2004). Multiple binding sites in fibrinogen for integrin alphaM beta2 (Mac-1). *J. Biol. Chem.* **279,** 44897–44966.

Liu, C. Y., Koehn, J. A., and Morgan, F. J. (1985). Characterization of fibrinogen New York 1. A dysfunctional fibrinogen with a deletion of B beta(9-72) corresponding exactly to exon 2 of the gene. *J. Biol. Chem.* **260,** 4390–4396.

Liu, J. N., Harpel, P. C., and Gurewich, V. (1994). Fibrin-bound lipoprotein(a) promotes plasminogen binding but inhibits fibrin degradation by plasmin. *Biochem.* **33,** 2554–2560.

Loewy, A. G., McDonagh, J., Mikkola, H., Teller, D. C., and Yee, V. C. (2000). Structure and function of Factor XIII. *In* ''Hemostasis and Thrombosis: Basic Principles and Clinical Practice'' (R. W. Colman, J. Hirsh, V. J. Marder, A. W. Clowes, and J. N. George, Eds.), pp. 233–247. Lippincott Williams & Wilkins, Philadelphia.

Lorand, L. (2001). Factor XIII: Structure, activation, and interactions with fibrinogen and fibrin. *Ann. NY Acad. Sci.* **936,** 291–311.

Lorand, L., Parameswaran, K. N., and Murthy, S. N. P. (1998). A double-headed Gly-Pro-Arg-Pro ligand mimics the functions of the E domain of fibrin for promoting

the end-to-end crosslinking of gamma chains by factor XIIIa. *Proc. Natl. Acad. Sci. USA* **95,** 1438–1442.

Loscalzo, J., Weinfield, M., Fless, G., and Scanu, A. M. (1990). Lipoprotein(a), fibrin binding, and plasminogen activation. *Atheroscler.* **10,** 240–245.

Louache, F., Debili, N., Cramer, E. M., Breton-Gorius, J., and Vainchenker, W. (1991). Fibrinogen is not synthesized by human megakaryocytes. *Blood* **77,** 311–316.

Lovely, R. S., Moaddel, M., and Farrell, D. H. (2003). Fibrinogen gamma′ chain binds thrombin exosite II. *J. Thromb. Haemost.* **1,** 124–131.

Ly, B., and Godal, H. C. (1973). Denaturation of fibrinogen: The protective effect of calcium. *Haemost.* **1,** 204–209.

Madrazo, J., Brown, J. H., Litvinovich, S., Dominguez, R., Yakovlev, S., Medved, L., and Cohen, C. (2001). Crystal structure of the central region of bovine fibrinogen (E5 fragment) at 1.4-A resolution. *Proc. Natl. Acad. Sci. USA* **98,** 11967–11972.

Makogonenko, E., Tsurupa, G., Ingham, K., and Medved, L. (2002). Interaction of fibrin(ogen) with fibronectin: Further characterization and localization of the fibronectin-binding site. *Biochem.* **41,** 7907–7913.

Marder, V. J., Shulman, N. R., and Carroll, W. R. (1969). High molecular weight derivatives of human fibrinogen produced by plasmin. *J. Biol. Chem.* **244,** 2111–2119.

Martinelli, R. A., and Scheraga, H. A. (1980). Steady-state kinetic study of the bovine thrombin-fibrinogen interaction. *Biochem.* **19,** 2343–2350.

Martinez, J., Keane, P. M., Gilman, P. B., and Palascak, J. E. (1983). The abnormal carbohydrate composition of the dysfibrinogenemia associated with liver disease. *Ann. NY Acad. Sci.* **408,** 388–396.

Martinez, J., Ferber, A., Bach, T. L., and Yaen, C. H. (2001). Interaction of fibrin with VE-cadherin. *Ann. NY Acad. Sci.* **936,** 386–405.

Matsuda, M., and Sugo, T. (2001). Hereditary disorders of fibrinogen. *Ann. NY Acad. Sci.* **936,** 65–68.

Matsuka, Y. V., Medved, L. V., Migliorini, M. M., and Ingham, K. C. (1996). Factor XIIIa-catalyzed cross-linking of recombinant alpha C fragments of human fibrinogen. *Biochem.* **35,** 5810–5816.

McDonagh, J. (2000). Dysfibrinogenemia and other disorders of fibrinogen structure or function. *In* ''Hemostasis and Thrombosis: Basic Principles and Clinical Practice'' (R. W. Colman, J. Hirsh, V. J. Marder, A. W. Clowes, and J. N. George, Eds.), pp. 855–892. Lippincott Williams & Wilkins, Philadelphia.

Medved, L., and Nieuwenhuizen, W. (2003). Molecular mechanisms of initiation of fibrinolysis by fibrin. *Thromb. Haemost.* **89,** 409–419.

Medved, L., Ugarova, T. P., Veklich, Y., Lukinova, N., and Weisel, J. W. (1990). Electron microscope investigation of the early stages of fibrin assembly. Twisted protofibrils and fibers. *J. Mol. Biol.* **216,** 503–509.

Medved, L., Litvinovich, S., Ugarova, T. P., Matsuka, Y. V., and Ingham, K. (1997). Domain structure and functional activity of the recombinant human fibrinogen gamma-module (gamma148–411). *Biochem.* **36,** 4685–4693.

Medved, L. V. (1990). Relationship between exons and domains in the fibrinogen molecule. *Blood Coag. Fibrinolysis* **1,** 439–442.

Medved, L. V., Litvinovich, S. V., and Privalov, P. L. (1986). Domain organization of the terminal parts in the fibrinogen molecule. *FEBS Lett.* **202,** 298–302.

Medved, L. V., Litvinovich, S. V., Ugarova, T. P., Lukinova, N. I., Kalikhevich, V. N., and Ardemasova, Z. A. (1993). Localization of a fibrin polymerization site complementary to Gly-His-Arg sequence. *FEBS Lett.* **320,** 239–242.

Meh, D. A., Siebenlist, K. R., Brennan, S. O., Holyst, T., and Mosesson, M. W. (2001). The amino acid sequence in fibrin responsible for high affinity thrombin binding. *Thromb. Haemost.* **85,** 470–474.

Mihalyi, E. (1988a). Clotting of bovine fibrinogen. Calcium binding to fibrin during clotting and its dependence on release of fibrinopeptide B. *Biochem.* **27,** 967–976.

Mihalyi, E. (1988b). Clotting of bovine fibrinogen. Kinetic analysis of the release of fibrinopeptides by thrombin and of the calcium uptake upon clotting at high fibrinogen concentrations. *Biochem.* **27,** 976–982.

Moaddel, M., Falls, L. A., and Farrell, D. H. (2000a). The role of gamma A/gamma′ fibrinogen in plasma factor XIII activation. *J. Biol. Chem.* **275,** 32135–32140.

Moaddel, M., Farrell, D. H., Daugherty, M. A., and Fried, M. G. (2000b). Interactions of human fibrinogens with factor XIII: Roles of calcium and the gamma′ peptide. *Biochem.* **39,** 6698–6705.

Moen, J. L., Gorkun, O. V., Weisel, J. W., and Lord, S. T. (2003). Recombinant BbetaArg14His fibrinogen implies participation of N-terminus of Bbeta chain in desA fibrin polymerization. *Blood* **102,** 2466–2471.

Mosesson, M. W. (2003). Antithrombin I. Inhibition of thrombin generation in plasma by fibrin formation. *Thromb. Haemost.* **89,** 9–12.

Mosesson, M. W. (2004a). Cross-linked gamma chains in a fibrin fibril are situated transversely between its strands: Yes. *J. Thromb. Haemost.* **2,** 388–393.

Mosesson, M. W. (2004b). Cross-linked gamma-chains in a fibrin fibril are situated transversely between its strands. *J. Thromb. Haemost.* **2,** 1467–1469.

Mosesson, M. W., Siebenlist, K. R., and Meh, D. A. (2001). The structure and biological features of fibrinogen and fibrin. *Ann. NY Acad. Sci.* **936,** 11–30.

Mosesson, M. W., Hainfeld, J., Wall, J., and Haschemeyer, R. H. (1981). Identification and mass analysis of human fibrinogen molecules and their domains by scanning transmission electron microscopy. *J. Mol. Biol.* **153,** 695–718.

Mosesson, M. W., DiOrio, J. P., Siebenlist, K. R., Wall, J. S., and Hainfeld, J. F. (1993). Evidence for a second type of fibril branch point in fibrin polymer networks, the trimolecular junction. *Blood* **82,** 1517–1521.

Mosesson, M. W., DiOrio, J. P., Müller, M. F., Shainoff, J. R., Siebenlist, K. R., Amrani, D. L., Homandberg, G. A., Soria, J., Soria, C., and Samama, M. (1987). Studies on the ultrastructure of fibrin lacking fibrinopeptide B (beta-fibrin). *Blood* **69,** 1073–1081.

Mullin, J. L., Gorkun, O. V., Binnie, C. G., and Lord, S. T. (2000). Recombinant fibrinogen studies reveal that thrombin specificity dictates order of fibrinopeptide release. *J. Biol. Chem.* **275,** 25239–25246.

Neerman-Arbez, M. (2001). Fibrinogen gene mutations accounting for congenital afibrinogenemia. *Ann. NY Acad. Sci.* **936,** 496–508.

Nham, S. U., and Fuller, G. M. (1986). Effect of fibrinogen degradation products on production of hepatocyte stimulating factor by a macrophage cell line (P388D1). *Thromb. Res.* **44,** 467–475.

Nieuwenhuizen, W. (2001). Fibrin-mediated plasminogen activation. *Ann. NY Acad. Sci.* **936,** 237–246.

Nossel, H. (1976). Radioimmunoassay of fibrinopeptides in relation to intravascular coagulation and thrombosis. *N. Engl. J. Med.* **295,** 428–432.

Odrljin, T. M., Rybarczyk, B. J., Francis, C. W., Lawrence, S. O., Hamaguchi, M., and Simpson-Haidaris, P. J. (1996). Calcium modulates plasmin cleavage of the fibrinogen D fragment gamma chain N-terminus: Mapping of monoclonal antibody J88B

to a plasmin sensitive domain of the gamma chain. *Biochim. Biophys. Acta* **1298,** 69–77.

Palumbo, J. S., Potter, J. M., Kaplan, L. S., Talmage, K., Jackson, D. G., and Degen, J. L. (2002). Spontaneous hematogenous and lymphatic metastasis, but not primary tumor growth or angiogenesis, is diminished in fibrinogen-deficient mice. *Cancer Res.* **62,** 6966–6972.

Pandya, B. V., Cierniewski, C. S., and Budzynski, A. Z. (1985). Conservation of human fibrinogen conformation after cleavage of the B beta chain NH2 terminus. *J. Biol. Chem.* **260,** 2994–3000.

Parrott, J. A., Whaley, P. D., and Skinner, M. K. (1993). Extrahepatic expression of fibrinogen by granulosa cells: Potential role in ovulation. *Endocrinology* **133,** 1645–1649.

Parry, D. A. D. (1978). Fibrinogen: A preliminary analysis of the amino acid sequences of the portions of the alpha, beta and gamma-chains postulated to form the interdomainal link between globular regions of the molecule. *J. Mol. Biol.* **120,** 545–551.

Pechik, I., Madrazo, J., Mosesson, M. W., Hernandez, I., Gilliland, G. L., and Medved, L. (2004). Crystal structure of the complex between thrombin and the central ''E'' region of fibrin. *Proc. Natl. Acad. Sci. USA* **101,** 2718–2723.

Plant, P. W., and Grieninger, G. (1986). Noncoordinate synthesis of the fibrinogen subunits in hepatocytes cultured under hormone-deficient conditions. *J. Biol. Chem.* **261,** 2331–2336.

Price, T. M., Strong, D. D., Rudee, M. L., and Doolittle, R. F. (1981). Shadow-cast electron microscopy of fibrinogen with antibody fragments bound to specific regions. *Proc. Natl. Acad. Sci. USA* **78,** 200–204.

Privalov, P. L., and Medved, L. V. (1982). Domains in the fibrinogen molecule. *J. Mol. Biol.* **159,** 665–683.

Profumo, A., Turci, M., Damonte, G., Ferri, F., Magatti, D., Cardinali, B., Cuniberti, C., and Rocco, M. (2003). Kinetics of fibrinopeptide release by thrombin as a function of CaCl2 concentration: Different susceptibility of FPA and FPB and evidence for a fibrinogen isoform-specific effect at physiological Ca2+ concentration. *Biochem.* **42,** 12335–12348.

Rand, M. D., Lock, J. B., van't Veer, C., Gaffney, D. P., and Mann, K. G. (1996). Blood clotting in minimally altered whole blood. *Blood* **88,** 3432–3445.

Rao, S., Poojary, M., Elliott, B., Melanson, L., Oriel, B., and Cohen, C. (1991). Fibrinogen structure in projection at 18 A resolution. Electron density by co-ordinated cryo-electron microscopy and X-ray crystallography. *J. Mol. Biol.* **222,** 89–98.

Remijn, J. A., Lounes, K. C., Hogan, K. A., Lord, S. T., Galanakis, D. K., Sixma, J. J., and De Groot, P. G. (2001). Reduced platelet adhesion in flowing blood to fibrinogen by alterations in segment gamma 316–322, part of the fibrin-specific region. *Ann. NY Acad. Sci.* **936,** 650–657.

Roberts, H. R., Stinchcombe, T. E., and Gabriel, D. A. (2001). The dysfibringenemias. *British J. Haematol.* **114,** 249–257.

Romer, J., Bugge, T. H., Pyke, C., Lund, L. R., Flick, M. J., Degen, J. L., and Dano, K. (1996). Impaired wound healing in mice with a disrupted plasminogen gene. *Nat. Med.* **2,** 287–292.

Rooney, M. M., Parise, L. V., and Lord, S. T. (1996). Dissecting clot retraction and platelet aggregation. Clot retraction does not require an intact fibrinogen gamma chain C terminus. *J. Biol. Chem.* **271,** 8553–8555.

Rooney, M. M., Farrell, D. H., van Hemel, B. M., de Groot, P. G., and Lord, S. T. (1998). The contribution of the three hypothesized integrin-binding sites in fibrinogen to platelet-mediated clot retraction. *Blood* **92,** 2374–2381.

Ruoslahti, E. (1988). Fibronectin and its receptors. *Annu. Rev. Biochem.* **57,** 375–413.

Ryan, E. A., Mockros, L. F., Weisel, J. W., and Lorand, L. (1999). Structural origins of fibrin clot rheology. *Biophys. J.* **77,** 2813–2826.

Sahni, A., and Francis, C. W. (2000). Vascular endothelial growth factor binds to fibrinogen and fibrin and stimulates endothelial cell proliferation. *Blood* **96,** 3772–3778.

Sahni, A., Odrljin, T. M., and Francis, C. W. (1998). Binding of basic fibroblast growth factor to fibrinogen and fibrin. *J. Biol. Chem.* **273,** 7554–7559.

Sahni, A., Sporn, L. A., and Francis, C. W. (1999). Potentiation of endothelial cell proliferation by fibrin(ogen)-bound fibroblast growth factor-2. *J. Biol. Chem.* **274,** 14936–14941.

Sahni, A., Guo, M., Sahni, S. K., and Francis, C. W. (2004). Interleukin-1beta but not IL-1alpha binds to fibrinogen and fibrin and has enhanced activity in the bound form. *Blood* **104,** 409–414.

Sakata, Y., and Aoki, N. (1980). Cross-linking of α2-plasmin inhibitor to fibrin by fibrin-stabilizing factor. *J. Clin. Invest.* **65,** 290–297.

Sakharov, D. V., and Rijken, D. C. (1995). Superficial accumulation of plasminogen during plasma clot lysis. *Circulation* **92,** 1883–1890.

Sakharov, D. V., Nagelkerke, J. F., and Rijken, D. C. (1996). Rearrangements of the fibrin network and spatial distribution of fibrinolytic components during plasma clot lysis. Study with confocal microscopy. *J. Biol. Chem.* **271,** 2133–2138.

Sakharov, D. V., Plow, E. F., and Rijken, D. C. (1997). On the mechanism of the antifibrinolytic activity of plasma carboxypeptidase B. *J. Biol. Chem.* **272,** 14477–14482.

Sasaki, T., Page, I. H., and Shainoff, J. R. (1966). Stable complex of fibrinogen and fibrin. *Science* **152,** 1069–1071.

Scrutton, M. C., Ross-Murphy, S. B., Bennett, G. M., Stirling, Y., and Meade, T. W. (1994). Changes in clot deformability – a possible explanation for the epidemiological association between plasma fibrinogen concentration and myocardial infarction. *Blood Coagul. Fibrinolysis* **5,** 719–723.

Shafer, J. A., and Higgins, D. L. (1988). Human fibrinogen. *CRC Critical Rev. in Clin. Lab. Sci.* **26,** 1–41.

Shah, G. A., Nair, C. H., and Dhall, D. P. (1987). Comparison of fibrin networks in plasma and fibrinogen solution. *Thromb. Res.* **45,** 257–264.

Shainoff, J. R., and Dardik, B. N. (1979). Fibrinopeptide B and aggregation of fibrinogen. *Science* **204,** 200–202.

Shainoff, J. R., and Dardik, B. N. (1983). Fibrinopeptide B in fibrin assembly and metabolism: Physiologic significance in delayed release of the peptide. *Ann. NY Acad. Sci.* **408,** 254–268.

Shainoff, J. R., Smejkal, G. B., Dibello, P. M., Chase, B., Mitkevich, O. V., and Lill, H. (2001). The fibrin intermediate, its place in the fibrinogen-fibrin transformation. *Ann. NY Acad. Sci.* **936,** 147–166.

Siebenlist, K. R., Meh, D. A., and Mosesson, M. W. (1996). Plasma factor XIII binds specifically to fibrinogen molecules containing gamma chains. *Biochem.* **35,** 10448–10453.

Siebenlist, K. R., DiOrio, J. P., Budzynski, A. Z., and Mosesson, M. W. (1990). The polymerization and thrombin-binding properties of des-(B beta 1–42)-fibrin. *J. Biol. Chem.* **265,** 18650–18655.

Spraggon, G., Everse, S. J., and Doolittle, R. F. (1997). Crystal structures of fragment D from human fibrinogen and its crosslinked counterpart from fibrin. *Nature* **389,** 455–462.

Standeven, K. F., Grant, P. J., Carter, A. M., Scheiner, T., Weisel, J. W., and Ariëns, R. A. S. (2003). Functional analysis of the fibrinogen AaThr312Ala polymorphism: Effects on fibrin structure and function. *Circulation* **107,** 2326–2330.

Stringer, H. A., and Pannekoek, H. (1995). The significance of fibrin binding by plasminogen activator inhibitor 1 for the mechanism of tissue-type plasminogen activator-mediated fibrinolysis. *J. Biol. Chem.* **270,** 11205–11208.

Strong, D. D., Laudano, A. P., Hawiger, J., and Doolittle, R. F. (1982). Isolation, characterization, and synthesis of peptides from human fibrinogen that block the staphylococcal clumping reaction and construction of a synthetic clumping particle. *Biochem.* **21,** 1414–1420.

Stryer, L., Cohen, C., and Langridge, R. (1963). Axial period of fibrinogen and fibrin. *Nature* **197,** 793–794.

Stubbs, M. T., and Bode, W. (1993). A model for the specificity of fibrinogen cleavage by thrombin. *Semin. Thromb. Hemost.* **19,** 344–351.

Suehiro, K., Gailit, J., and Plow, E. F. (1997). Fibrinogen is a ligand for integrin alpha5beta1 on endothelial cells. *J. Biol. Chem.* **272,** 5360–5366.

Suenson, E., Lützen, O., and Thorsen, S. (1984). Initial plasmin-degradation of fibrin as the basis of a positive feedback mechanism in fibrinolysis. *Eur. J. Biochem.* **149,** 193–200.

Suh, T. T., Holmback, K., Jensen, N. J., Daugherty, C. C., Small, K., Simon, D. I., Potter, S., and Degen, J. L. (1995). Resolution of spontaneous bleeding events but failure of pregnancy in fibrinogen-deficient mice. *Genes Develop.* **9,** 2020–2033.

Takeda, Y. (1966). Studies of the metabolism and distribution of fibrinogen in healthy men with autologous 125-I-labeled fibrinogen. *J. Clin. Invest.* **45,** 103–111.

Thorsen, S. (1992). The mechanism of plasminogen activation and the variability of the fibrin effector during tissue-type plasminogen activator-mediated fibrinolysis. *Ann. NY Acad. Sci.* **667,** 52–63.

Tooney, N. M., and Cohen, C. (1972). Microcrystals of a modified fibrinogen. *Nature* **237,** 23–25.

Tooney, N. M., and Cohen, C. (1974). Crystallisation of a modified fibrinogen. *Nature* **251,** 659–660.

Tooney, N. M., and Weisel, J. W. (1979). Fibrinogen: Structure and assembly. *In* ''Fibrous Proteins: Scientific, Industrial and Medical Aspects'' (D. A. D. Parry, and L. K. Creamer, Eds.), Vol. 1, pp. 429–454. Academic Press, London.

Torbet, J. (1986). Fibrin assembly in human plasma and fibrinogen/albumin mixtures. *Biochem.* **25,** 5309–5314.

Torbet, J. (1995). The thrombin activation pathway modulates the assembly, structure, and lysis of human plasma clots *in vitro. Thromb. Haemost.* **73,** 785–792.

Townsend, R. R., Heller, D. N., Fenselau, C. C., and Lee, Y. C. (1984). Determination of the sialylation pattern of human fibrinogen glycopeptides with fast atom bombardment. *Biochem.* **23,** 6389–6392.

Townsend, R. R., Hilliker, E., Li, Y. T., Laine, R. A., Bell, W. R., and Lee, Y. C. (1982). Carbohydrate structure of human fibrinogen. Use of 300-MHz 1H-NMR to characterize glycosidase-treated glycopeptides. *J. Biol. Chem.* **257,** 9704–9710.

Tran, H., Tanaka, A., Litvinovich, S. V., Medved, L. V., Haudenschild, C. C., and Argraves, W. S. (1995). The interaction of fibulin-1 with fibrinogen. A potential role in hemostasis and thrombosis. *J. Biol. Chem.* **270,** 19458–19464.

Tsurupa, G., and Medved, L. (2001). Fibrinogen alpha C domains contain cryptic plasminogen and tPA binding sites. *Ann. NY Acad. Sci.* **936,** 328–330.

Tsurupa, G., Ho-Tin-Noe, B., Angles-Cano, E., and Medved, L. (2003). Identification and characterization of novel lysine-independent apolipoprotein(a)-binding sites in fibrin(ogen) alphaC-domains. *J. Biol. Chem.* **278,** 37154–37159.

Tuszynski, G. P., Srivastava, S., Switalska, H. I., Holt, J. C., Cierniewski, C. S., and Niewiarowski, S. (1985). The interaction of human platelet thrombospondin with fibrinogen. *J. Biol. Chem.* **260,** 12240–12245.

Ugarova, T. P., Solovjov, D. A., Zhang, L., Loukinov, D. I., Yee, V. C., Medved, L. V., and Plow, E. F. (1998). Identification of a novel recognition sequence for integrin alphaM beta2 within the gamma-chain of fibrinogen. *J. Biol. Chem.* **273,** 22519–22527.

Ugarova, T. P., Lishko, V. K., Podolnikova, N. P., Okumura, N., Merkulov, S. M., Yakubenko, V. P., Yee, V. C., Lord, S. T., and Haas, T. A. (2003). Sequence gamma 377–395(P2), but not gamma 190–202(P1), is the binding site for the alpha MI-domain of integrin alpha M beta 2 in the gamma C-domain of fibrinogen. *Biochem.* **42,** 9365–9373.

van Zonneveld, A. J., Veerman, H., and Pannekoek, H. (1986). On the interaction of the finger and the kringle-2 domain of tissue-type plasminogen activator with fibrin. Inhibition of kringle-2 binding to fibrin by epsilon-amino caproic acid. *J. Biol. Chem.* **261,** 14214–14218.

Veklich, Y. I., Gorkun, O. V., Medved, L. V., Nieuwenhuizen, W., and Weisel, J. W. (1993). Carboxyl terminal portions of the alpha chains of fibrinogen and fibrin: Localization by electron microscopy and the effects of isolated alphaC fragments on polymerization. *J. Biol. Chem.* **268,** 13577–13585.

Vindigni, A., and Di Cera, E. (1996). Release of fibrinopeptides by the slow and fast forms of thrombin. *Biochem.* **35,** 4417–4426.

Wang, W., Boffa, M. B., Bajzar, L., Walker, J. B., and Nesheim, M. E. (1998). A study of the mechanism of inhibition of fibrinolysis by activated thrombin-activable fibrinolysis inhibitor. *J. Biol. Chem.* **273,** 27176–27181.

Wangh, L. J., Holland, L. J., Spolski, R. J., Aprison, B. S., and Weisel, J. W. (1983). Xenopus fibrinogen. Characterization of subunits and hormonal regulation of biosynthesis. *J. Biol. Chem.* **258,** 4599–4605.

Weisel, J. W. (1986a). The electron microscope band pattern of human fibrin: Various stains, lateral order, and carbohydrate localization. *J. Ultrastruct. Mol. Struct. Res.* **96,** 176–188.

Weisel, J. W. (1986b). Fibrin assembly. Lateral aggregation and the role of the two pairs of fibrinopeptides. *Biophys. J.* **50,** 1079–1093.

Weisel, J. W. (2004a). Cross-linked gamma chains in a fibrin fibril are situated transversely between its strands: No. *J. Thromb. Haemost.* **2,** 394–399.

Weisel, J. W. (2004b). The mechanical properties of fibrin for basic scientists and clinicians. *Biophys. Chem.* **112,** 267–276.

Weisel, J. W. (2004c). Cross-linked gamma-chains in a fibrin fibril are situated transversely between its strands. *J. Thromb. Haemost.* **2,** 1467–1469.

Weisel, J. W., and Nagaswami, C. (1992). Computer modeling of fibrin polymerization kinetics correlated with electron microscope and turbidity observations: Clot structure and assembly are kinetically controlled. *Biophys. J.* **63,** 111–128.

Weisel, J. W., and Cederholm-Williams, S. A. (1997). Fibrinogen and fibrin: Characterization, processing and medical applications. *In* ''Handbook of Biodegradable Polymers'' (A. J. Domb, J. Kost, and D. M. Wiseman, Eds.), pp. 347–365. Harwood, Amsterdam.

Weisel, J. W., and Medved, L. (2001). The structure and function of the αC domains of fibrinogen. *Ann. NY Acad. Sci.* **936,** 312–327.

Weisel, J. W., Warren, S. G., and Cohen, C. (1978). Crystals of modified fibrinogen: Size, shape and packing of molecules. *J. Mol. Biol.* **126,** 159–183.

Weisel, J. W., Phillips, G. J., and Cohen, C. (1981). A model from electron microscopy for the molecular structure of fibrinogen and fibrin. *Nature* **289,** 263–267.

Weisel, J. W., Phillips, G. J., and Cohen, C. (1983). The structure of fibrinogen and fibrin: II. Architecture of the fibrin clot. *Ann. NY Acad. Sci.* **408,** 367–379.

Weisel, J. W., Nagaswami, C., and Makowski, L. (1987). Twisting of fibrin fibers limits their radial growth. *Proc. Natl. Acad. Sci. USA* **84,** 8991–8995.

Weisel, J. W., Stauffacher, C. V., Bullitt, E., and Cohen, C. (1985). A model for fibrinogen: Domains and sequence. *Science* **230,** 1388–1391.

Weisel, J. W., Nagaswami, C., Vilaire, G., and Bennett, J. S. (1992). Examination of the platelet membrane glycoprotein IIb-IIIa complex and its interation with fibrinogen and other ligands by electron microscopy. *J. Biol. Chem.* **267,** 16637–16643.

Weisel, J. W., Nagaswami, C., Korsholm, B., Petersen, L. C., and Suenson, E. (1994). Interactions of plasminogen and polymerizing fibrin and its derivatives monitored with a photoaffinity cross-linker and electron microscopy. *J. Mol. Biol.* **235,** 1117–1135.

Weisel, J. W., Nagaswami, C., Woodhead, J. W., Higazi, A. A.-R., Cain, W. J., Marcovina, S. M., Koschinsky, M. L., Cines, D. B., and Bdeir, K. (2001). The structure of lipoprotein(a) and ligand-induced conformational changes. *Biochem.* **40,** 10424–10435.

Williams, R. C. (1983). Morphology of fibrinogen monomers and of fibrin protofibrils. *Ann. NY Acad. Sci.* **408,** 180–193.

Wiman, B., and Collen, D. (1978). Molecular mechanism of physiological fibrinolysis. *Nature* **272,** 549–550.

Wolfenstein, T. C., and Mosesson, M. W. (1981). Carboxy-terminal amino acid sequence of a human fibrinogen gamma-chain variant (gamma′). *Biochem.* **20,** 6146–6149.

Woodhead, J. L., Nagaswami, C., Matsuda, M., Arocha-Piñango, C. L., and Weisel, J. W. (1996). The ultrastructure of fibrinogen Caracas II molecules, fibers and clots. *J. Biol. Chem.* **271,** 4946–4953.

Xu, W., Chung, D. W., and Davie, E. W. (1996). The assembly of human fibrinogen. The role of the amino-terminal and coiled-coil regions of the three chains in the formation of the alphagamma and betagamma heterodimers and alphabetagamma half-molecules. *J. Biol. Chem.* **271,** 27948–27953.

Xu, X., and Doolittle, R. F. (1990). Presence of a vertebrate fibrinogen-like sequence in an echinoderm. *Proc. Natl. Acad. Sci. USA* **87,** 2097–2101.

Yakovlev, S., Makogonenko, E., Kurochkina, N., Nieuwenhuizen, W., Ingham, K., and Medved, L. (2000). Conversion of fibrinogen to fibrin: Mechanism of exposure of tPA- and plasminogen-binding sites. *Biochem.* **39,** 15730–15741.

Yamazumi, K., and Doolittle, R. F. (1992). The synthetic peptide Gly-Pro-Arg-Pro-amide limits the plasmic digestion of fibrinogen in the same fashion as calcium ion. *Prot. Sci.* **1,** 1719–1720.

Yang, Z., Pandi, L., and Doolittle, R. F. (2002). The crystal structure of fragment double-D from cross-linked lamprey fibrin reveals isopeptide linkages across an unexpected D-D interface. *Biochem.* **41,** 15610–15617.

Yang, Z., Kollman, J. M., Pandi, L., and Doolittle, R. F. (2001). Crystal structure of native chicken fibrinogen at 2.7 A resolution. *Biochemistry* **40,** 12515–12523.

Yee, V. C., Pedersen, L. C., Le Trong, I., Bishop, P. D., Stenkamp, R. E., and Teller, D. C. (1994). Three-dimensional structure of a transglutaminase: Human blood coagulation factor XIII. *Proc. Natl. Acad. Sci. USA* **91,** 7296–7300.

Yee, V. C., Pratt, K. P., Cote, H. C., Trong, I. L., Chung, D. W., Davie, E. W., Stenkamp, R. E., and Teller, D. C. (1997). Crystal structure of a 30 kDa C-terminal fragment from the gamma chain of human fibrinogen. *Structure* **5,** 125–138.

Yoshida, N., Wada, H., Morita, K., Hirata, H., Matsuda, M., Yamazumi, K., Asakura, S., and Shirakawa, S. (1991). A new congenital abnormal fibrinogen Ise characterized by the replacement of B beta glycine-15 by cysteine. *Blood* **77,** 1858–1863.

Yu, S., Sher, B., Kudryk, B., and Redman, C. M. (1984). Fibrinogen precursors. Order of assembly of fibrinogen chains. *J. Biol. Chem.* **259,** 10574–10581.

Zhang, J. Z., and Redman, C. M. (1996). Fibrinogen assembly and secretion. Role of intrachain disulfide loops. *J. Biol. Chem.* **95,** 30083–30088.

MOLECULAR STRUCTURE OF THE COLLAGEN TRIPLE HELIX

By BARBARA BRODSKY AND ANTON V. PERSIKOV

Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

I. Introduction 302
II. The Biological Role and Occurrence of the Collagen Triple-Helix Motif in Proteins 303
III. Molecular Structure: Collagen and Collagen Model Peptides 307
A. Sequence Dependence of Triple Helix Twist 308
B. Hydrogen Bonding 311
C. Hydration Networks 314
D. Side Chain Interactions 315
E. Molecular Packing 316
F. Molecular Dynamics 318
IV. Mechanism of Hydroxyproline Stabilization 319
V. Breaks in the Gly-X-Y Repeating Pattern 323
VI. Amino Acid Sequence and Stability 323
VII. Ligand Binding 326
VIII. Mutations and Disease 329
IX. Conclusions 332
X. Abbreviations and Notation 332
References 333

Abstract

The molecular conformation of the collagen triple helix confers strict amino acid sequence constraints, requiring a $(\text{Gly-X-Y})_n$ repeating pattern and a high content of imino acids. The increasing family of collagens and proteins with collagenous domains shows the collagen triple helix to be a basic motif adaptable to a range of proteins and functions. Its rodlike domain has the potential for various modes of self-association and the capacity to bind receptors, other proteins, GAGs, and nucleic acids. High-resolution crystal structures obtained for collagen model peptides confirm the supercoiled triple helix conformation, and provide new information on hydrogen bonding patterns, hydration, sidechain interactions, and ligand binding. For several peptides, the helix twist was found to be sequence dependent, and such variation in helix twist may serve as recognition features or to orient the triple helix for binding. Mutations in the collagen triple-helix domain lead to a variety of human disorders. The most common mutations are single-base substitutions that lead to the

DOI: 10.1016/S0065-3233(04)70009-1
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

replacement of one Gly residue, breaking the Gly-X-Y repeating pattern. A single Gly substitution destabilizes the triple helix through a local disruption in hydrogen bonding and produces a discontinuity in the register of the helix. Molecular information about the collagen triple helix and the effect of mutations will lead to a better understanding of function and pathology.

I. Introduction

Collagen has always been considered to be an important protein because of its abundance in the human body and its commercial applications. Fifty years ago, the molecular conformation for collagen was proposed to be a novel triple-helix structure. Fiber diffraction analysis and model building, together with early amino acid composition and sequence data, led to the concept of three chains, each in a polyproline-II-like conformation, supercoiled about a common axis (Ramachandran, 1967; Ramachandran and Kartha, 1955; Rich and Crick, 1955, 1961). The close packing of the three chains near the common axis places steric constraints on every third position, such that only glycine can be accommodated without chain distortion. This generates the $(\text{Gly-X-Y})_n$ repeating sequence, recognized as the signature of a collagen. Residues in the X and Y positions are largely solvent accessible and can accommodate any side-chain. In fact, imino acids are particularly favorable because their fixed ψ angle and restricted φ angle are close to those found in the triple helix. When Pro is incorporated into the Y position in a collagen chain, it becomes posttranslationally modified to hydroxyproline (Hyp), which has a highly stabilizing effect on the triple helix.

Early on, it was realized that peptides could be useful models for the features of collagen. Polyglycine and polyproline played an important role in elucidating the triple-helix structure (Cowan *et al.*, 1955; Rich and Crick, 1955), and polytripeptides were designed to adopt the collagen conformation. Initially, the peptides were heterogeneous polymers, but with the advent of solid state peptide synthesis, peptides of defined length and sequence could be synthesized to clarify principles of triple-helix stability and mimic biological activity (Fields and Prockop, 1996; Goodman *et al.*, 1998; Heidemann and Roth, 1982; Jenkins and Raines, 2002). Most peptides are studied as monomers that self-associate into trimers, while, in some cases, crosslinked trimers are synthesized to facilitate trimer formation, stabilize the triple helix, or create heterotrimers. Triple-helical peptides are proving more amenable to structural characterization than the longer native collagen molecules, and peptide investigations have led to clarification of the molecular details of collagen structure by

X-ray crystallography, nuclear magnetic resonance (NMR), and other biophysical techniques.

Because of its strict structural constraints, triple-helix domains can be identified on the basis of amino acid sequence. With many amino acid sequences of proteins available through molecular biology, there is clearly a growing family of collagen proteins with a common triple-helix motif present in diverse tissue structures and an expanding collection of collagen-like domains in various other proteins. The collagen triple helix can form a straight or kinked rod, has the capacity for self-association into various supramolecular structures, and has the potential to bind ligands and receptors. These structural features of collagen have adapted to a wide range of biological roles, and now the triple-helix motif is more appropriately viewed in a general context as a basic and versatile protein motif. Additionally, a growing number of mutations in collagens and collagen-like domains have been associated with specific diseases. The need to understand the effects of these mutations has lent a growing impetus to characterization of the normal and variant forms of molecular conformation and higher order structure.

This review will focus on recent advances in the molecular structure of the collagen triple helix, including information from high-resolution structures, the effect of amino acid sequence on stability, ligand binding, and disease. Important articles on the molecular properties of collagen were published in earlier volumes of this series (Harrington and von Hippel, 1961; Privalov, 1982; Traub and Piez, 1971). There are excellent reviews of earlier work on the molecular structure of collagen (Fraser and MacRae, 1973; Fraser *et al.*, 1987; Ramachandran, 1967) and several comprehensive overviews of collagens, their structures, assembly, and biochemistry (Kielty and Grant, 2002; Myllyharju and Kivirikko, 2004).

II. The Biological Role and Occurrence of the Collagen Triple-Helix Motif in Proteins

Collagens are defined as structural molecules in the extracellular matrix that contain a triple-helix domain; at this time, 27 distinct types of human collagens have been identified (Fig. 1; Table I; Kielty and Grant, 2002; Myllyharju and Kivirikko, 2004). These different collagen types carry out specialized functions in diverse tissues and have distinctive modes of supramolecular organization. Some molecules are homotrimers, while others are heterotrimers, with two or three distinguishable chain types. The most abundant of these are found in characteristic collagen fibrils (major types I, II, III, and minor types V and XI), which form the structural basis of skin, tendon, bone, cartilage, and other tissues. Other

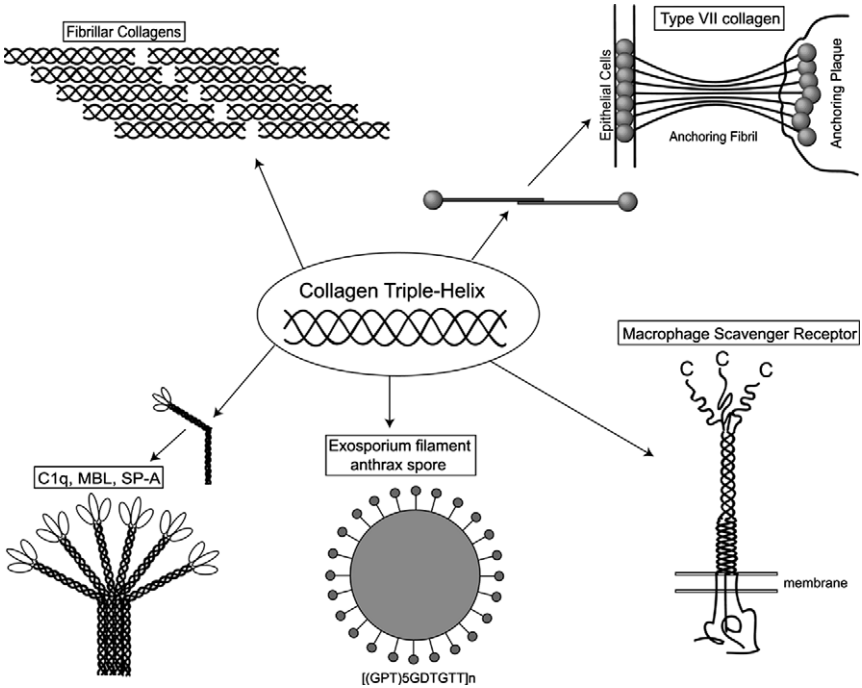


FIG. 1. Illustration of some of the biological forms of the collagen triple-helix domains.

collagens are found on the surface of fibrils (FACIT types IX, XII, XIV, XVI, XIX, XX, XXI, XXII, XVI); within basement membrane networks (type IV collagen); in hexagonal networks (types VIII and X); as beaded filaments (type VI); in the anchoring fibrils of skin (type VII); or as membrane proteins (types XIII, XVII, XXIII, XXV). Type XVIII collagen is a basement membrane heparan sulfate proteoglycan, and contains a C-terminal noncollagenous fragment, endostatin, which inhibits angiogenesis and tumor growth (Kawashima *et al.*, 2003). A thorough discussion of the various collagen types, their structures, and their functions is included in several excellent reviews (Kielty and Grant, 2002; Myllyharju and Kivirikko, 2004).

In addition to being the defining feature of collagens, the collagen triple helix is present as a motif in a variety of proteins, many of which are involved in host-defense functions (Fig. 1; Table I; Kielty and Grant, 2002; Lu *et al.*, 2002; Matsuzawa *et al.*, 2004; Myllyharju and Kivirikko, 2004;

TABLE I
Occurrence of Proteins with Collagen Triple-Helix Domains

Proteins	Supramol	Tissue	Number of $(GXY)_n$	Breaks
Vertebrates				
Type I collagen	D-periodic fibril	Tendon, bone, skin	338	0
Type II collagen	D-periodic fibril	Cartilage, vitreous	338	0
Type IV collagen	Network	Basement membrane	437	21
Type VII collagen	Antiparallel	Anchoring fibrils	472	20
Type IX collagen	Surface of fibrils	Ubiquitous	85	4
Type XII collagen	(FACIT)		197	5
C1q	Hexamer	Serum, complement	24	1
MBP	Oligomer	Serum	19	1
SP-A	Hexamer	Lung surfactant	23	1
SP-D	Tetramer	Lung surfactant	59	0
MSR	Membrane	Macrophage	24	0
Invertebrates				
C. elegans cuticle collagen		Exoskeleton	~9, ~40	1–4
Drosophila type IV collagen	Network	Basement membrane	444	17
Hydra minicollagen	Disulfide linked polymers	Nematocyst inner wall	12–16	0
Bacteria and Viruses (~120 types)				
Scl1	Cell surface	*Streptococcus* group A	50	0
BclA	Exosporium filament	*Bacillus anthracis*	70	0
Lymphocytis disease virus	Unknown	Unknown	55	1
Shrimp white spot virus	Unknown	Unknown	389	1

Shirai *et al.*, 1999). The kinked triple-helix domain in C1q binds to the serine proteases C1r and C1s, and mediates the self-association of six trimer molecules that generates the bouquet-type structure important for activity. The family of collectins, which includes the mannose-binding lectin SP-A and SP-D, all contain a collagenous domain, a coiled coil α-helical domain, and a terminal carbohydrate recognition domain. Ficolins have a triple-helix domain with a terminal carbohydrate binding fibringen domain. Both collectins and ficolins are host-defense molecules, which bind to the carbohydrate groups of microbes, leading to complement activation and phagocytosis (Lu *et al.*, 2002). The triple-helix domain of the macrophage scavenger receptor is responsible for ligand recognition (Doi *et al.*, 1993; Shirai *et al.*, 1999), while the collagenous tail of the asymmetric form of acetylcholinesterase binds to heparan sulfate, localizing the enzyme to the neuromuscular junction (DePrez *et al.*, 2000). In addition to these examples, numerous other proteins have been observed to contain collagen-like domains, some of which are listed in Table I.

Orthologues of fibril-forming collagens and of type IV collagen are found in a number of invertebrates, in addition to specialized collagens such as the cuticle collagens of *C. elegans* and the hydra nematocyst minicollagens (Exposito *et al.*, 2002). It was thought that collagen was a defining feature of multicellular animals, but recent observations show triple-helix domains present in bacteria and viruses. Collagen-like domains Scl1 and Scl2 were observed in proteins expressed on the cell surface of group A streptococcus, and were shown to adopt a triple-helix conformation (Xu *et al.*, 2002). A highly repetitive collagen-like domain was also identified in the filament protein of the exosporium of anthrax spores, and the length of this collagenous domain appears to determine the filament length (Sylvestre *et al.*, 2003). A search of the genomes of various bacteria and viruses indicated the presence of at least 100 novel proteins containing collagen-related structural motifs (Rasmussen *et al.*, 2003). These findings confirm that the collagen triple-helix motif can be part of many different kinds of proteins and can fill a wider than expected set of biological niches (Table I).

The occurrence of the collagen triple helix illustrates its role in protein function (Fig. 1). In all cases, the triple helix forms a rodlike structure, but this rod can have a kink, as in C1q and MBL, or have flexible interruptions, as in type IV collagen. In most cases, the collagen triple helix self-associates to form a higher-order structure. Many possible modes of association are seen, including staggered parallel (fibrils), parallel unstaggered (stalk of MBP, SP-A and C1q), antiparallel (type VII and type VI collagens) and other more complex relationships in networks and fibers. A number of collagens and proteins with collagenous domains contain

transmembrane domains (MSR, and collagen types XIII, XVII, XXIII, and XXV). Fibrillar collagen structure and network-forming collagen structures are the focus of Chapters 10 and 11. The triple helix is an important ligand binding domain in most proteins, binding to receptors, proteases, extracellular matrix proteins, GAGs, nucleic acids, and lipids in different cases (see section VII below). In contrast to early thinking, it is increasingly clear that the collagen triple helix is not a domain responsible for trimerization. The triple helix folds very slowly, and is usually dependent on a neighboring coiled coil or globular domain for nucleation (McAlinden *et al.*, 2003). But once trimerized, its rodlike structure has the potential for a wide range of modes of self-association and the ability to bind diverse ligands, as seen in the growing collection of proteins with triple-helix domains.

III. Molecular Structure: Collagen and Collagen Model Peptides

The advances in fiber diffraction analysis, amino acid composition data, molecular modeling, and polypeptide structures all converged during the mid-1950s to clarify the unique structure of collagen. The supercoiled triple-helix conformation was proposed for collagen in 1955 independently by Ramachandran, who proposed a similar model without supercoiling in 1954, (Ramachandran and Kartha, 1954); by Rich and Crick (1955); and by Cowan, McGavin, and North (Cowan *et al.*, 1955). In this conformation, three polypeptide chains, each in an extended left-handed polyproline II-helix conformation, are supercoiled in a right-handed manner around a common axis, with a stagger of one residue between adjacent chains. The most accurate model available for collagen is based on linked-atom least-squares refinement of the Rich and Crick II model using the excellent fiber diffraction data from highly stretched partially dehydrated kangaroo tail tendon, as reported by Fraser *et al.* (1979). The basic conformation was confirmed and complemented by a range of spectroscopic and hydrodynamic studies on collagens from various species. The historical developments and collagen investigations are reviewed in Fraser and MacRae (1973).

Although there were early reports of crystallization of a cyanogen bromide fragment of collagen (Yonath and Traub, 1975) and important NMR studies were carried out on isotopically labeled collagen in tissues (Batchelder *et al.*, 1982; Sarkar *et al.*, 1983, 1987), the collagen molecule itself has not proved amenable to investigations at the molecular level. The path to the molecular details of the collagen triple helix has been through collagen model peptides, which have yielded high-resolution X-ray structures and allowed NMR characterization of dynamic and conformational features. At

this time, the crystal structures of nine different triple-helical peptides have been reported, one with a bound integrin domain (Table II; Fig. 2). The fiber diffraction models of collagen left unresolved controversies about the nature of interchain hydrogen bonding, including the possibility of hydrogen bonds involving alpha carbon atoms, and the precise geometrical parameters of both the basic helix and the supercoil. The high-resolution crystal structures of peptides confirmed the conformation derived from fiber diffraction data of collagen, and resolved a number of long-standing controversies about hydrogen bonding and hydration. The molecular details also raise new issues for consideration, including variability in helix twist and the mechanism of hydroxyproline stabilization.

A. *Sequence Dependence of Triple Helix Twist*

Fiber diffraction patterns of collagen in tail tendon are indexed with a 10/3 symmetry (10 units in 3 turns) (Fraser *et al.*, 1979; Rich and Crick, 1961). Therefore, it was surprising when Okuyama *et al.* (1981) reported that the first crystal structure of a collagen-like peptide, $(\text{Pro-Pro-Gly})_{10}$, had a 7/2 symmetry (7 units in 2 turns). The difference in triple-helix symmetry between the $(\text{Pro-Pro-Gly})_{10}$ crystal and collagen fibers could have arisen from crystal packing effects, the unusually high imino acid content in the peptide compared with collagen, or from the absence of Hyp in the peptide. Recent crystallographic structures have confirmed the 7/2 symmetry is present in $(\text{Gly-Pro-Hyp})_{10}$ and the G→A peptide, as well as $(\text{Pro-Pro-Gly})_{10}$ (Table II). The EKG peptide, which is homologous to $(\text{Pro-Hyp-Gly})_{10}$ but with one Glu-Lys-Gly triplet in the center, also shows strict 7/2 symmetry (Kramer *et al.*, 2000).

However, a nonuniform helical twist has been observed in the crystal structures of peptides T3-785 and IBP (Emsley *et al.*, 2000, 2004; Kramer *et al.*, 1999, 2001). These two peptides each can be viewed as having three zones: N-terminal Gly-Pro-Hyp repeats; a central collagen sequence; and C-terminal Gly-Pro-Hyp repeats (Fig. 3). In T3-785, the two terminal Gly-Pro-Hyp regions show 7/2 symmetry, while the central Gly-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala region is closer to 10/3 symmetry. In IBP, the terminal Gly-Pro-Hyp repeats have 7/2 symmetry, and the central Gly-Phe-Hyp-Gly-Glu-Arg sequence is intermediate between 7/2 and 10/3. In each case, these three zones are slightly bent and twisted with respect to each other. It appears that the 7/2 symmetry is generated by the steric restrictions of repeating Gly-imino acid-imino acid units and is maintained when only one Gly-X-Y triplet is introduced (e.g., EKG peptide). The presence of 2 or 3 tripeptide units, where X and Y are not imino acids, starts to change the helix twist towards 10/3 symmetry.

TABLE II
Collagen-Like Peptides with High-Resolution Crystal Structures Entered in the Protein Data Bank

Peptide	Sequence	Number of residues	PDB ID	Resolution (Å)	Helical twist	Reference
G->A	$(POG)_4$ POA $(POG)_5$	30	1CAG	1.85	7/2	Bella *et al.*, 1994
T3-785	$(POG)_3$ITGARGLAG$(POG)_4$	30	1BKV	2.00	Variable	Kramer *et al.*, 1999
EKG	$(POG)_4$ EKG $(POG)_5$	30	1QSU	1.75	7/2	Kramer *et al.*, 2000
Hyp-	$(POG)_4$ PG $(POG)_5$	29	1EI8	2.00	7/2	Liu, 2000
PPG10	$(PPG)_{10}$	30	1K6F	1.30	7/2	Berisio *et al.*, 2002
GPP-foldon	$(GPP)_9$-foldon	27	1NAY	2.60	7/2	Stetefeld *et al.*, 2003
PPG9	$(PPG)_9$	27	1ITT	1.00	7/2	Hongo *et al.*, unpub.
POG10	$(POG)_{10}$	30	1V4F	1.26	7/2	Okuyama *et al.*, unpub.
Integrin binding peptide	$(GPO)_2$ GFOGER $(GPO)_3$	21	1Q7D	2.10	Variable	Emsley *et al.*, 2004
	In complex with integrin	21	1DZI	1.80	Variable	Emsley *et al.* 2000

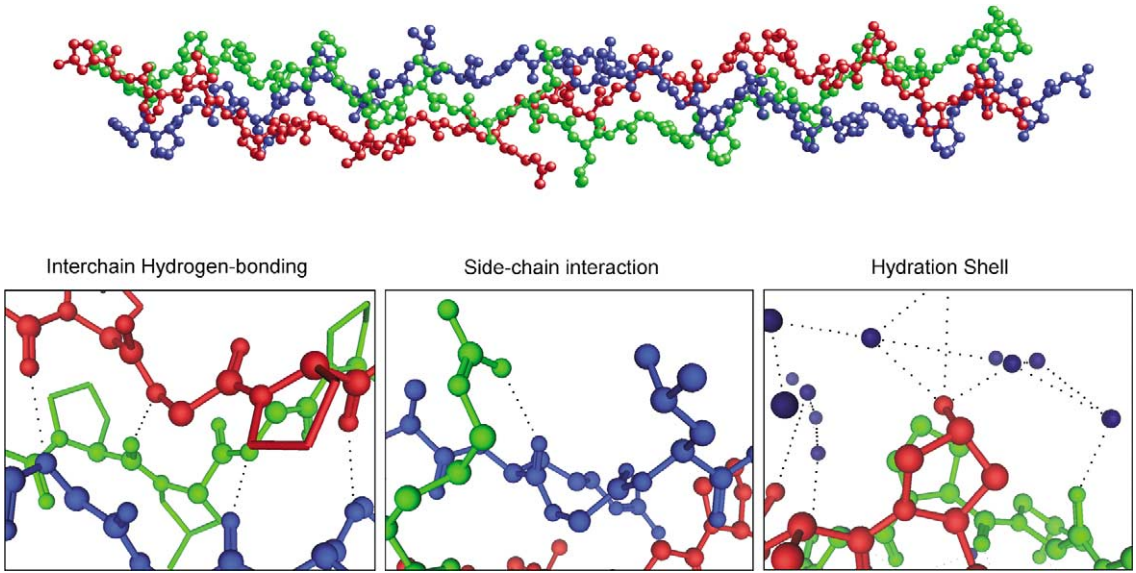


FIG. 2. Molecular structures of peptide T3-785 showing hydrogen bonding, sidechain orientation, and hydration.

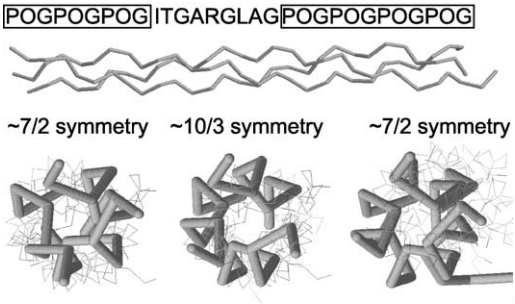


FIG. 3. Variation in helix twist seen in the three domains of peptide T3-785.

These peptide results suggest the collagen molecule, with its varied Gly-X-Y sequence, is likely to have a nonuniform helical twist along its length. It is likely sequences poor in imino acids will have a symmetry close to 10/3, while stretches of Gly-Pro-Hyp units may have 7/2 symmetry. In type I collagen, the longest sequence of repeating Gly-Pro-Hyp triplets is found at the C-terminus, and 7/2 symmetry could play a role in its functioning as a nucleation domain (Xu *et al.*, 2003). It is not clear at this time whether collagen chains have a continuous variation in superhelix twist, or domains of different symmetry connected with kinks and bends. The X-ray pattern of tail tendons does show a predominance of the 10/3 symmetry, but coherent regions may be overrepresented in fiber diffraction patterns. The conformational differences between the 7-fold and 10-fold symmetric triple-helices are subtle, but the small difference in units per turn (3.50 vs. 3.33) would translate into a significant difference in the helix repeat over a long distance and could affect recognition.

B. Hydrogen Bonding

Hydrogen bonding is a critical part of triple-helix stabilization. The very favorable enthalpy reported for collagen compared to other proteins is consistent with the importance of hydrogen bonding (Privalov, 1982). The triple helix has repetitive backbone hydrogen bonding networks, but differs from beta sheets or alpha helices in that the repeating tripeptide unit consists of three nonequivalent peptide groups, and not all backbone peptide groups participate in hydrogen bonding. It was clear from earliest models that one strong interchain peptide NH...OC bond could be formed per Gly-X-Y tripeptide unit (Ramachandran, 1967; Rich and Crick,

1961). Ramachandran originally argued that a second NH...OC bond was possible when the residue in the X position was not Pro (Ramachandran and Kartha, 1955), but because of distortion and steric problems, this view was modified to suggest an interaction mediated by water (Ramachandran and Chandrasekharan, 1968).

All crystal structures show a hydrogen bond between the NH of Gly in one chain and the C=O of the residue in the X position of the neighboring chain, as predicted (Fig. 4). In addition, peptides with sequences where the X position is occupied by a residue other than Pro, show a second interchain hydrogen bond between the amide group of the X position residue and the C=O of the Gly residue, which is mediated by one water molecule (Emsley *et al.*, 2000, 2004; Kramer *et al.*, 1999, 2000). This second set of hydrogen bonds connects chains in a direction opposite to that of the first set, as proposed by Ramachandran and Chandrasekharan (1968). In some cases, the water molecules involved in the NH (X position)...CO (Gly) hydrogen bond make additional hydrogen bonds with Hyp or side chains. When the Y position is occupied by an amino acid, rather than an imino acid, its amide group is hydrated by water molecules directed into the solvent, which is not likely to contribute to stability.

The hydrogen bonding patterns seen in peptide crystals are consistent with hydrogen exchange studies and thermodynamic analyses (Fan *et al.*, 1993; Privalov, 1982; Yee *et al.*, 1974). NMR hydrogen exchange studies on 15N-labeled Gly-Leu-Ala residues in the central zone of peptide T3-785 demonstrated that Gly NH exchanged the slowest, Leu NH exchanged almost as slowly, and Ala NH showed an exchange rate similar to peptides exposed to bulk solvent (Fan *et al.*, 1993). The water-mediated hydrogen bond of Leu, involving the X position amide, slows the hydrogen exchange process almost as much as the direct Gly interchain hydrogen bonds. Hydrogen exchange studies on collagen also show two distinct slowly exchanging sets of amide hydrogens (Privalov, 1982; Yee *et al.*, 1974). Therefore, the second set of water-mediated hydrogen bonds is likely to contribute to stability and may serve to reinforce the triple helix in regions lacking Pro.

The possibility of C^{α}H...CO hydrogen bonds in polyglycine II and collagen was raised by Ramachandran (1967) and Krimm and Kuroiwa (1968). A detailed analysis of the high resolution structure of the G→A peptide supports the existence of two kinds of such bonds: an H-bond between the Gly C^{α} in one chain and the Gly and Pro C=O groups from the other two chains; and an H-bond connecting Hyp C^{α} in one chain with a Pro C=O group on the neighboring chain (Bella and Berman, 1996). These C^{α}H...CO bonds are suggested to define a network of weak interactions that may add additional stability to the strong NH...CO bonds.

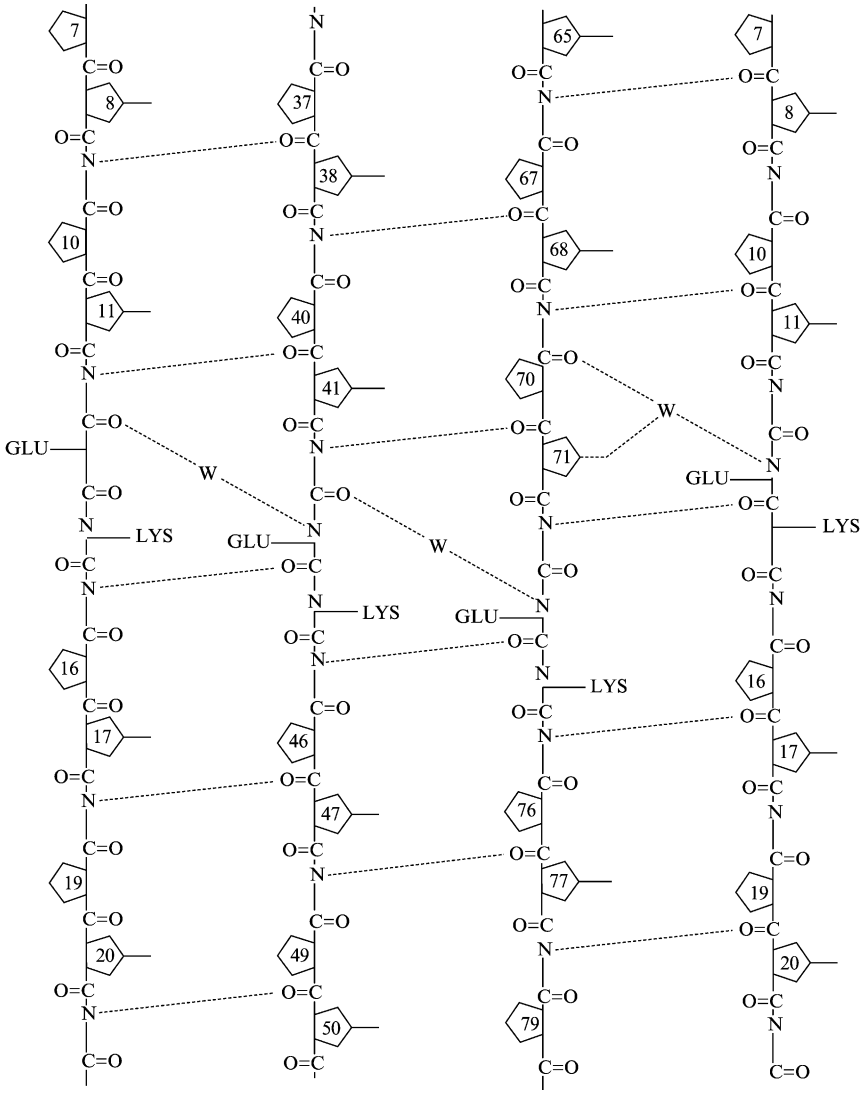


FIG. 4. Schematic of the hydrogen bonding patterns seen in the high-resolution structure of EKG peptide.

C. *Hydration Networks*

Water is an integral part of the collagen molecule (Fraser and MacRae, 1973; Harrington and Von Hippel, 1961; Traub and Piez, 1971). The collagen triple helix has tightly bound water, and experimental evidence supports it being a highly ordered hydration network (Berendsen and Migchelsen, 1965; Grigera and Berendsen, 1979; Suzuki *et al.*, 1980). The water was proposed to bind to available backbone carbonyls; since two C=O groups per tripeptide do not participate direct NH. . .CO bonds, and original concepts of single water binding were expanded to suggest waters bonded to each other and then to several main chain atoms (Ramachandran and Chandrasekharan, 1968). The function of hydroxyproline has also been related to this hydration network. On the basis of an elegant analysis of thermodynamics and hydrogen exchange data, Privalov found that Hyp stabilization of collagen is correlated with increased enthalpy (Privalov, 1982). Since Hyp cannot directly bond with the backbone of the molecule, enthaplic stabilization supports the involvement of an ordered water hydrogen-bonding network.

The determination of the high-resolution structure of the G→A peptide provided the first visualization of the elaborate water network that surrounds collagen molecules (Fig. 2; Bella *et al.*, 1994, 1995). Water molecules are seen to bridge C=O groups within and between molecules and to link C=O and Hyp hydroxyl groups within and between molecules. Often four or five water molecules participate in these bridges, with pentagonal, clathrate-like geometry. Strikingly repetitive networks of these water patterns are seen along the chain. Water tends to order near hydrophobic surfaces, and Bella *et al.* (1995) suggest the abundance of pentagonal water clusters in the triple-helical cylinder of hydration may result from the solvent exposure of nonpolar Pro and Hyp residues, together with the availability of backbone CO groups and Hyp OH groups to anchor the ordered water network to the peptide. Initially, questions were raised about whether these intricate water networks were real (Engel and Prockop, 1998), but the increasing number of high-resolution structures have confirmed that extended water networks are an inherent feature of all collagen triple-helix peptide crystal structures (Berisio *et al.*, 2001, 2002; Kramer *et al.*, 2000, 2001). NMR studies have indicated the kinetically labile nature of this collagen hydration shell (Melacini *et al.*, 2000).

The nature and order of the water network appears to depend on the molecular packing in the crystal and the specific sequence present. In peptide T3-785, less ordered water was seen in the terminal Gly-Pro-Hyp regions that were not well packed, than in the central Gly-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala region with close molecular packing (Kramer *et al.*, 2001).

For the EKG peptide, there was some disruption of the regular hydration pattern in the central charged region, where the packing is also less dense (Kramer *et al.*, 2000). In addition to interactions involving backbone CO groups and Hyp, water is also seen to interact with sidechains, mediating hydrogen bonds between ionizable sidechains and between sidechains and the backbone (Kramer *et al*, 2000, 2001). In the central region of peptide T3-785, where there are no Hyp residues present, the waters involved in the hydrogen bonds between NH(X)...CO(Gly) also interact with Thr and Arg in the Y positions rather than Hyp (Kramer *et al.*, 2001).

The high-resolution structures of triple-helical peptides allow visualization of the ordered water network that was expected from studies on collagen, and has elucidated the specific bonds involved in this network and their repetitive nature. There is strong evidence supporting the role of this ordered water in maintaining molecular packing in fibrils (Leikin *et al.*, 1994, 1995, 1997). The biological and physical significance of this water network with respect to molecular stability has been a subject of much debate (Jenkins and Raines, 2002). The Hyp hydroxyl groups are clearly key players in repetitive hydration networks in the high-resolution structures, but one cannot assess from a crystal structure alone whether the entropic cost of localizing the waters will be less or greater than the highly favorable enthalpy from hydrogen bond formation in the collagen molecule (see Section IV below).

D. Side Chain Interactions

The availability of high-resolution structures of peptides EKG, T3-785, and IBP, which include residues other than Pro and Hyp in the X and Y positions, offers the opportunity to investigate the conformation and interactions of sidechains from residues typically found within the collagen triple helix. In the peptide with an EKG tripeptide sequence, the Lys and Glu residues did not form direct intermolecular or intramolecular ion pairs, even though such pairs are sterically feasible (Kramer *et al.*, 2000). Instead, the Lys side chains bond to Y position carbonyl groups of an adjacent chain, while one Glu directly interacts with a Hyp hydroxyl group. There was also a range of water-mediated interactions involving the polar sidechains. In peptide T3-785, with the central region Gly-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala, the Arg side chains make direct contacts with backbone carbonyl groups on an adjacent chain, confirming predictions of this interaction and clarifying the high stability of Arg in the Y position (Fig. 2; Kramer *et al.*, 2001; Vitagliano *et al.*, 1993). In T3-785, the Arg sidechains also make hydrophobic interactions with Leu and Ile sidechains

from the same or neighboring molecules, forming nonpolar clusters that minimize exposure to solvent. The Thr sidechains are involved in bonding with the water, which mediates hydrogen bonds between the amide groups from the X position and Gly C═O groups. The participation of Thr in the Y position in the water network, much as Hyp does when it is in the Y position, suggests Thr could play a similar stabilizing role in invertebrates and bacteria. In IBP, with the central Gly-Phe-Hyp-Gly-Glu-Arg sequence, one Glu is involved in an intrahelix water mediated interaction with Arg, while Glu (as well as the Arg sidechains) are involved in direct interactions with backbone peptide groups of neighboring molecules.

All sidechains in X and Y positions of the triple helix are exposed to solvent and appear to have multiple options of interacting through solvent and with available backbone carbonyl groups, in addition to sidechain-sidechain interactions. NMR studies on collagen fibrils show there are two interconverting conformations of Leu in fibrils, and considerable reorientation around the helix axis (Batchelder *et al.*, 1982; Sarkar *et al.*, 1983). This suggests there may be switching between alternative interaction sets in solution and even in fibrils. Thus, the sidechain orientations and interactions seen in the crystal structure may represent one of a number of possibilities, rather than a uniquely determined interaction.

E. Molecular Packing

One surprise in the X-ray diffraction studies was the resemblance of the molecular packing of peptides in the crystal structure to that seen by fiber diffraction for collagen molecules in tendon fibrils. The packing of molecules in the crystals of G→A, EKG, and POG is quasi-hexagonal with intermolecular spacings near 14Å, an arrangement and distance very similar to collagen molecular packing (Fig. 5; Fraser *et al.*, 1983). The 14Å spacing between triple-helices is spanned by highly ordered water molecules that connect neighboring molecules. This type of network is consistent with the attractive hydration forces that have been suggested to function in collagen assemblies (Leikin *et al.*, 1994, 1995, 1997). The strong similarity between peptide and collagen organization suggests that the 14Å lateral spacing and quasi-hexagonal packing observed in collagen assemblies are sequence-independent and determined by the effective diameters of the cylinder of hydration coating the triple helix.

The most common intermolecular interaction—observed in GEK, GPO, and T3-785 crystals—was Hyp-Hyp hydrogen bonds between neighboring triple-helices (Fig. 5; Berisio *et al.*, 2001; Kramer *et al.*, 2000, 2001). The observation of Hyp-Hyp intermolecular interactions in the peptide crystals and the report that recombinant unhydroxylated collagen will not form

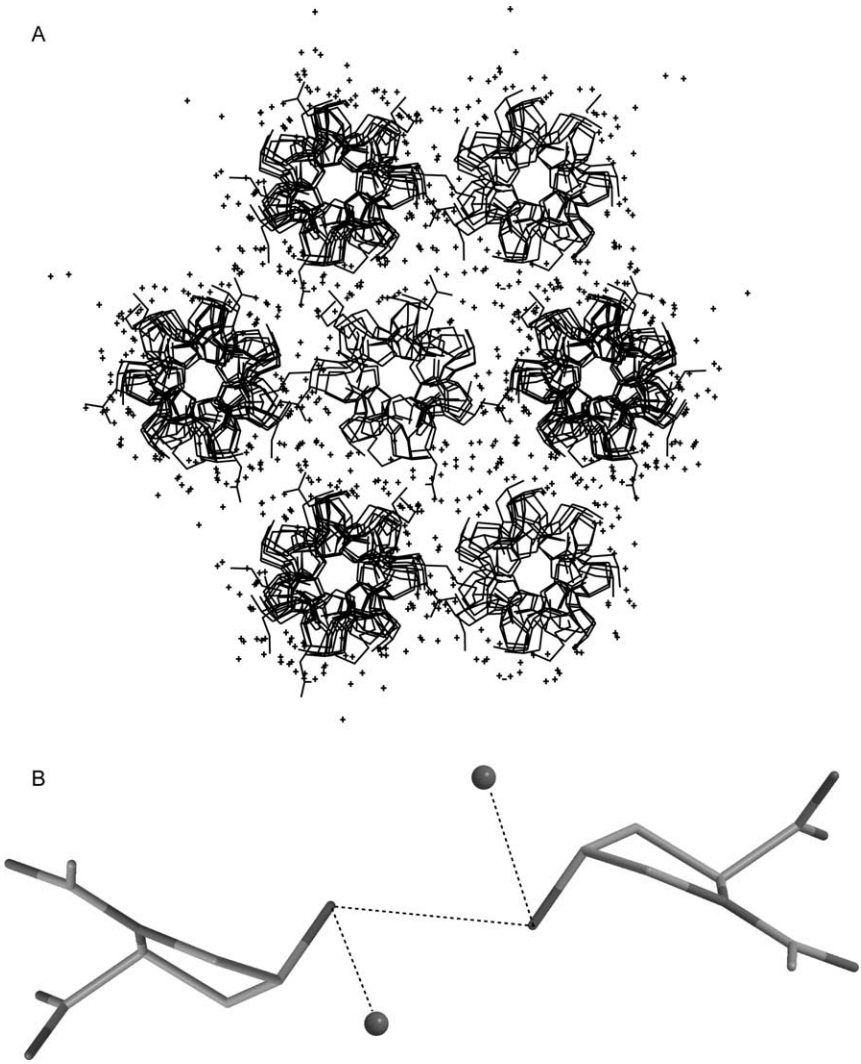


FIG. 5. (A) Cross-section of the molecular packing of the EKG peptide in the crystal. (B) The Hyp-Hyp interactions seen between neighboring molecules in the crystal.

fibrils (Perret *et al.*, 2001) support the original hypothesis of Gustavson (1955) that Hyp plays an important role in molecular association and fibril formation. There are Hyp residues all along triple helices, and direct interactions between two Hyp groups in adjacent molecules could be important in stabilization or orientation of molecular packing.

While the peptide lateral packing in crystals resembles that in collagen fibrils, the axial relationships, such as the D staggering (234 residues or 670Å), are likely to reflect interactions between charged and/or hydrophobic side chains. It was assumed that such interactions would be direct, between side chains of neighboring molecules, but in the crystal structures, an Arg or Lys group of one molecule usually binds to the backbone of a neighboring molecule (directly or through water) rather than to another side chain (Emsley *et al.*, 2004; Kramer *et al.*, 2001). The EKG peptide shows a staggering of neighboring molecules with charged Lys and Glu residues of one peptide aligned with the charged N and C-termini of other molecules (Kramer *et al.*, 2000), supporting the importance of electrostatic interactions in determining axial relationships between molecules.

F. Molecular Dynamics

Even though the collagen triple helix is considered a rigid rod, NMR, fluorescence, and molecular dynamics studies indicate there are sequence-dependent motions of the backbone and side chains. It appears that Gly-Pro-Hyp is the most rigid, as well as the most stabilizing, sequence. In model peptides, all backbone amides show high-order parameters typical of a rigid structure, but the Gly amide shows a slower hydrogen exchange rate when in a Gly-Pro-Hyp environment (Gly-Pro-Hyp-**Gly**-Pro-Hyp–Gly-Pro-Hyp) than in the imino acid poor environment (Gly-Ala-Arg-**Gly**-Leu-Ala-Gly-Pro-Hyp) (Fan *et al.*, 1993). Side chains of residues other than Pro and Hyp have considerable mobility in the collagen molecule, usually adopting more than one preferred orientation (Batchelder *et al.* 1982). Fluorescence studies of host-guest peptides showed that Trp in the Y position has a more restricted motion than Trp in the X position, which is consistent with the greater solvent accessibility of the X position (Simon-Lukasik *et al.*, 2003). Solid-state NMR studies of collagen with isotopically labeled residues indicate that the imino acid rings have angular fluctuations and that amino acid side chains are dynamic, reorienting around at least two side chain bonds (Sarkar *et al.*, 1983, 1987). Significant azimuthal motions around the helix axis were also observed (Sarkar *et al.*, 1983).

The sequence-dependent mobility of the collagen triple helix is important for its biological function. The unique collagenase cleavage site of types I and III collagen is at a boundary with a stable N-terminal region and a highly unstable C-terminal sequence (Fields, 1991). A molecular dynamics study carried out on the collagenase cleavage site of type III collagen, based on the crystal structure of peptide T3-785, suggested an alternative partially unfolded conformation that could expose the triple-helix backbone to cleavage (Stultz, 2002). Recently, experimental studies

indicate that collagenase binds and locally unwinds the triple-helix conformation in type I collagen prior to hydrolysis of its peptide bonds (Chung *et al.*, 2004). Locally unstable regions, which would be expected to be more dynamic, have also been implicated in collagen binding of various ligands as discussed below in Section VII.

Computational studies starting with collagen triple-helix structures now available in the Protein Data Bank are becoming increasingly important for collagen. As described above, molecular dynamics simulations can suggest models that are amenable to experimental verification. Computational analyses and molecular dynamics studies have also been carried out on Gly substitutions in the triple helix (see Section VIII below).

IV. Mechanism of Hydroxyproline Stabilization

Typically, the imino acid content of collagen is about 20%, with at least half in the form of hydroxyproline. Since collagens are the only animal proteins with a high content of hydroxyproline, an important role is indicated for this posttranslationally modified residue. Early proposals by Gustavson supported a role for Hyp in the stabilization of collagen fibrils, while later studies focused on a role in stabilization of the triple-helical molecule (Fraser and MacRae, 1973; Gustavson, 1955; Ramachandran, 1967). Evidence that Hyp contributes to triple-helix stability came from the greater stability of (Pro-Hyp-Gly)$_{10}$ (T_m=60 °C) compared with (Pro-Pro-Gly)$_{10}$ (T_m=30 °C), and from the decreased stability of unhydroxylated collagen synthesized in the presence of prolyl hydroxylase inhibitors (Rosenbloom *et al.*, 1973; Sakikabara *et al.*, 1973). Recent studies on recombinant collagen expressed in tobacco plants confirm the decrease in collagen stability when hydroxylation of proline is absent (Perret *et al.*, 2001).

The nature and the basis of the stabilizing effect of hydroxyproline have been actively investigated through studies on repeating polytripeptides and host-guest peptides (Table III). The predominant hydroxyproline found in collagen domains is 4-Hyp (i.e., on the gamma carbon of the imide ring). 4-Hyp is exclusively in the R diastereoisomer form and the importance of this stereospecificity is illustrated by the inability of (Pro-4SHyp-Gly)$_{10}$ to form a triple helix (Inouye *et al.*, 1976). The location of (4R)-Hyp in the Y position of the Gly-X-Y repeating sequence is also critical since (4RHyp-Pro-Gly)$_{10}$, with Hyp in the X position rather than the Y, does not adopt a triple-helical conformation (Inouye *et al.*, 1982). Interestingly, triple-helices can be formed with (4R)-Hyp in the X position if (4R)-Hyp or Thr is in the Y position (Table III; Bann and Bachinger, 2000; Berisio *et al.*, 2004). A very small amount of (3S)-Hyp is present in most

TABLE III
Thermal Stability of Repeating Polytripeptides $(Gly\text{-}X\text{-}Y)_{10}$ and Host-Guest Peptides, $(GPO)_3$-GXY-$(GPO)_4$, Containing Hydroxyproline and Fluoroproline*

Triplet	T_m (°C)		
	Repeating	Host-guest	References
Pro-Pro-Gly	33 (Persikov *et al.*, 2003)	45.5	(Persikov *et al.*, 2003)
Pro-(4R)Hyp-Gly	60 (Persikov *et al.*, 2003)	47.3	(Persikov *et al.*, 2003)
(4R)Hyp-Pro-Gly	<4	43.0	(Inouye *et al.*, 1982; Persikov *et al.*, 2003)
Pro-(4S)Hyp-Gly	<4		Inouye *et al.*, 1982
(4S)Hyp-Pro-Gly	<4		Inouye *et al.*, 1982
(4R)Hyp-(4R)Hyp-Gly	65	47.3	(Berisio *et al.*, 2004; Persikov *et al.*, 2003)
Gly-Pro-(3S)Hyp	<4	37.5	(Mizuno *et al.*, 2004)
Gly-(3S)Hyp-(4R)Hyp	<4	49.6	(Mizuno *et al.*, 2004)
Pro-(4R)Flp-Gly	87	43.7	(Holmgren *et al.*, 1999; Persikov *et al.*, 2003)
(4R)Flp-Pro-Gly	<4		(Doi *et al.*, 2003)
Pro-(4S)Flp-Gly	<4		(Bretscher *et al.*, 2001)
(4S)Flp-Pro-Gly	58		(Doi *et al.*, 2003)
Gly-Pro-Thr	<4		(Bann and Bachinger, 2000)
Gly-(4R)Hyp-Thr	19		(Bann and Bachinger, 2000)

* Hodges and Raines (2003) carried out experiments for (4R)Flp and (4S)Flp in the X position using a $(X\text{-}Y\text{-}Gly)_7$ design, and reached similar conclusions to those seen for the ten repeating units.

collagens in the X position, but model polytripeptides, including this residue in the X or Y position, were not triple-helical and it had a destabilizing influence in host-guest peptides (Jenkins *et al.*, 2003; Mizuno *et al.*, 2004).

The mechanism of stabilization by Hyp has generated considerable controversy (Fig. 6). Since the hydroxyl group of Hyp is directed outside and is unable to participate in any direct hydrogen bonds within the triple helix, the possibility of hydrogen bonding through water was suggested (Ramachandran *et al.*, 1973; Suzuki *et al.*, 1980). The strongest evidence for Hyp stabilization through water-mediated H bonding comes from Privalov's analysis showing that, in different collagens, increased Hyp content is correlated with increased enthalpic stabilization (Privalov,

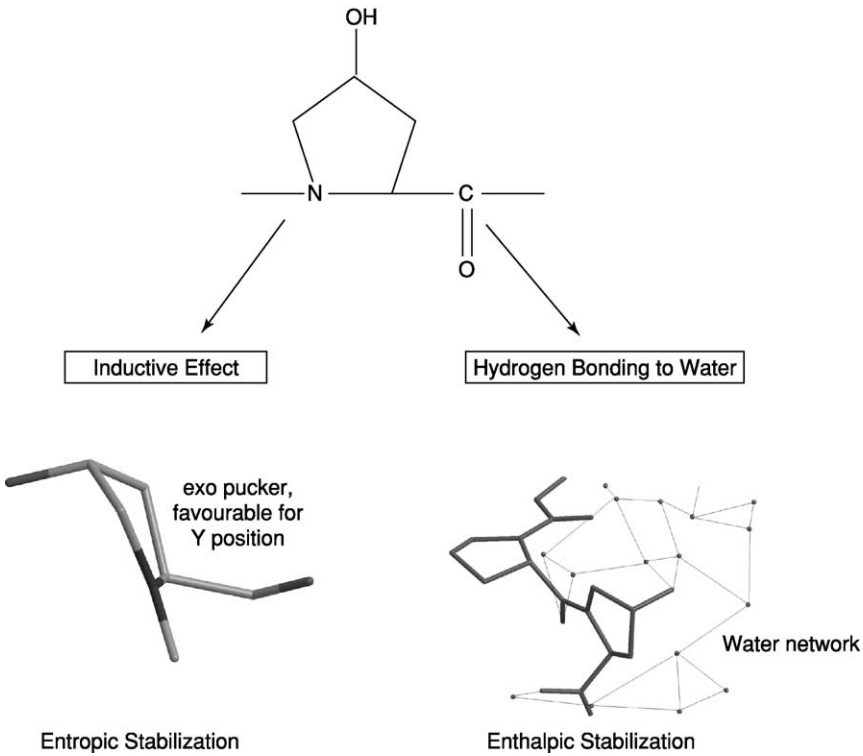


FIG. 6. Two proposed mechanisms for hydroxyproline stabilization.

1982). When the first high-resolution crystal structures were solved, hydroxyl groups of Hyp were seen to act as anchoring points for both intramolecular and intermolecular multispan water bridges (Bella *et al.*, 1994, 1995). It was suggested that the upward pucker of Hyp rings could be important because of Hyp's favorable orientation for binding to water networks. Arguments were raised about the existence of this hydration network in the crystal structures (Engel and Prockop, 1998), but such organized water was later confirmed in various laboratories for all triple-helical peptides. A highly ordered network was also seen in (Pro-Pro-Gly)$_{10}$ crystals, in the absence of Hyp, and questions were raised about whether the increased stability of (Pro-Hyp-Gly)$_{10}$ peptide originates from the hydrogen bonding of hydroxyprolines with the hydration network (Holmgren *et al.*, 1999).

In the late 1990s, Ronald Raines's group stimulated this area by taking a new approach, incorporating fluoroproline (Flp) residues in the design of collagen-like peptides (Jenkins and Raines, 2002). Fluorine has a greater electron-withdrawing effect on the imide ring than a hydroxyl group, and has a low tendency to form hydrogen bonds. The observation that (Pro-Flp-Gly)$_{10}$ (T_m~90 °C) was much more stable than (Pro-Hyp-Gly)$_{10}$ (T_m~60 °C) led to investigations about the basis of Flp stabilization and a reexamination of the nature of Hyp stabilization. Raines proposed that electron-withdrawing effects in (4R)-Flp would lead to a strong preference for the puckering of the imino ring in an up (or exo) conformation, and to the favoring of the trans vs. cis imide peptide bond (De Rider *et al.*, 2002). The puckering of imidic acids in the triple helix seen in the crystal structures largely confirmed the presence of down or endo for imino acids in the X position and up or exo conformation for imino acids in the Y position, as previously concluded from fiber diffraction models (Fraser *et al.*, 1979) and NMR data (Fan *et al.* 1993). Vitagliano *et al.* (2001) suggested that the stabilization of the triple helix by (4R)-Hyp in the Y position is due to its preference for the exo puckered conformation. Observations on the stabilizing effect of 4SFlp, which strongly prefers the down pucker in the X position, is consistent with this hypothesis (Table III; Doi *et al.*, 2003; Hodges and Raines, 2003). Hydrogen bonding through a water network is not a possibility for Flp, so its stabilization of the triple helix, which is even greater than that of Hyp, must come only from ring pucker and trans/cis preferences, mediated by its electronegative, inductive effect.

Thus, recent experiments involving Flp as well as Hyp, in different stereoisomeric forms and in X and Y positions, have made major contributions to our understanding of triple-helix stabilization (Berisio *et al.*, 2002, 2004; Persikov *et al.*, 2003). However, it is not clear that Flp and Hyp stabilize the collagen triple helix by the same mechanism. The experimental results support a predominant role for the inductive effect and its subsequent effect on the exo ring pucker in the stabilization of the triple helix by Flp. However, the large enthalpic stabilization seen for (Gly-Pro-Hyp)$_{10}$ and for collagens suggests that hydrogen bonding of Hyp through hydration networks is likely to play an important role in hydroxyproline stabilization, in addition to its favoring of the exo pucker.

The unique Hyp residue may be important both at the molecular and higher-order structure levels in collagen. Bacteria and viruses lack prolyl hydroxylase and have no hydroxyproline in their collagen-like domains. The significant differences in their amino acid composition and sequence compared to animal collagens suggest the use of alternative stabilization strategies for the triple helix (Rasmussen *et al.*, 2003).

V. Breaks in the Gly-X-Y Repeating Pattern

The basic amino acid features of the collagen triple helix can be reconsidered in light of the many sequences now available for this motif. Early studies on fibril-forming type I collagen indicated that Gly should be every third residue throughout a triple-helix domain. But it is clear that the sequences of most collagens and collagen-like domains include sporadic interruptions in the repeating pattern, and that the fibril-forming collagens, with their $(\text{Gly-X-Y})_{338\text{-}341}$ structure, are unusual in having such a strict requirement. Type IV and VII collagen each contain more than 20 breaks in their very long triple-helical sequences. A number of collagens have between 2–10 interruptions (e.g., types VIII, X, VI), while some domains have a single break (e.g., mannose binding protein, SP-A, C1q) (Table I). The structural and functional implications of these breaks in the Gly-X-Y repeat may vary in different molecules. Early studies on C1q indicated that the site of the single interruption in the triple helix led to a rigid kink (Kilchherr *et al.*, 1985), while some breaks in type IV collagen and type XVI have been associated with flexible sites (Kassner *et al.*, 2004; Siebold *et al.*, 1987). As seen in the G→A peptide structure (where one Gly is replaced by an Ala) or the Hyp–peptide structure (where one Hyp is absent), one consequence of a break is to induce a discontinuity in registration between the domains on either side, serving to mark different regions (see Section VIII below; Bella *et al.*, 1994; Liu, 2000). A strict requirement for Gly as every third residue must be met in a domain of at least 6–7 tripeptide units in order to adopt a triple-helical conformation, but this conformation may be extended through breaks that do not meet this requirement.

VI. Amino Acid Sequence and Stability

All collagens have a high content of stabilizing imino acids, including the posttranslationally formed hydroxyproline, but there is a range of these important residues in collagens and collagen-like domains. In using translated sequences of DNA, it is assumed, based on experimental studies of collagens C1q and MBL, that all Pro residues in the Y position will be hydroxylated to Hyp, since prolyl hydroxylase specifically recognizes this position. Type I collagen has an equal amount of Pro and Hyp, while most collagens have more Hyp than Pro residues. The thermal stability of collagens from different animals correlates with their upper environmental temperature and with their imidic acid and Hyp content (Burjanadze, 2000; Rigby and Robinson, 1975).

Sequences of the form Gly-Pro-Hyp confer maximal stability to the collagen triple helix, while variations in the identities of the residues in the X and Y positions determine global thermal stability and modulate local stability and energetics that are required for self-association, recognition, and binding. Recent studies on recombinant collagen have shown that there are domains of varying stability along the collagen chain (Steplewski *et al.*, 2004). As seen in Section IV, the use of repeating polytripeptides have been extremely important in defining conformational features, but repeats of most tripeptide sequences in collagen will not lead to a stable triple helix. One approach to analysis of more varied collagen sequences has been host-guest peptides, where one or two tripeptide units are introduced into a constant, stabilizing Gly-Pro-Hyp framework. This approach has been used to determine the propensity of all 20 amino acids for the X position in a Gly-X-Hyp triplet, and for all 20 residues in the Y position in a Gly-Pro-Y triplet (Persikov *et al.*, 2000) The most stabilizing residues for the X position are Pro, Glu, Ala, Lys, Arg, Gln, and Asp, while the most stabilizing residues for the Y position are Hyp, Arg, Met, Ile, Gln, and Ala. The least stabilizing residues for both positions are the aromatic residues and Gly. These data provide a scale for the propensities of a given residue in the X or Y position in a homotrimer, measured in terms of the destabilizing influence of each residue compared to Gly-Pro-Hyp.

The propensities of individual residues for the triple-helix conformation may be affected by intramolecular interactions. For a given Gly-X1-Y1-Gly-X2-Y2 sequence, interchain interactions are sterically possible. Because the three chains (A, B, and C) are staggered by one residue, there are close contacts between sidechain X1 of chain A and Y1 of chain B, and between Y1 of chain C and X2 of chain A. In addition, intrachain X1–X2 and Y1–Y2 interactions are possible. Measurements of thermal stability of host-guest peptides with varying residues in the guest positions show that a subset of all sterically possible ion pairs lead to favorable interactions, and some favorable hydrophobic interactions as well (Persikov *et al.*, 2002, 2005). Helix stability is increased by interchain interactions such as Gly-Arg-Asp and Gly-Lys-Asp sequences, but the most energetically favorable interactions occur when sequences of the form Lys-Gly-Glu and Lys-Gly-Asp are present. The thermal stability of the 400 possible Gly-X-Y sequences are presented in Table IV, with the experimentally determined values shown in bold; those values predicted on the basis of the additivity of individual residues in the X and Y positions are shown in italics. A significant variation in thermal stability can be seen varying from the most stabilizing Gly-Pro-Hyp unit ($T_m = 47\,^\circ C$) to the low stability Gly-Gly-Phe ($T_m = 20\,^\circ C$).

TABLE IV

Predicted (*italics*) and Experimentally Observed (*bold*) T_m Values (°C) for all Possible Gly-X-Y Tripeptide Units in a Triple-Helix, Based on Host-Guest Peptide Studies*

X\Y	O	R	M	I	Q	A	V	E	T	C	K	H	S	D	G	L	N	Y	F	W
P	**47**	**47**	**43**	**42**	**41**	**41**	**40**	**40**	**40**	**38**	**37**	**36**	**35**	**34**	**33**	**32**	**30**	**30**	**28**	**26**
E	**43**	**40**	*38*	*37*	**38**	**35**	**35**	*35*	**36**	*33*	**35**	*31*	*31*	**30**	*29*	*28*	**30**	*26*	*24*	*22*
A	**42**	**38**	*37*	*36*	*36*	**33**	*34*	*34*	*34*	*32*	**31**	*30*	**33**	**33**	*27*	**28**	*26*	*25*	**22**	*21*
K	**42**	**39**	*37*	*36*	**39**	*35*	*34*	**35**	*34*	*32*	*31*	*30*	*29*	**36**	*27*	*27*	**32**	*24*	*23*	*20*
R	**41**	*41*	*36*	*35*	*35*	*34*	*33*	**34**	*33*	*31*	**30**	*29*	**31**	**35**	*26*	*26*	*25*	*24*	*22*	*19*
Q	**40**	**40**	*36*	*35*	*34*	*34*	*33*	*33*	*33*	*31*	**33**	*29*	*28*	*27*	*26*	*26*	*25*	*23*	*22*	*19*
D	**40**	**37**	*35*	*34*	*34*	**32**	*33*	*33*	*33*	*31*	**31**	*29*	*28*	*27*	*26*	*26*	*25*	*23*	*21*	*19*
L	**39**	*39*	*34*	*33*	**36**	**31**	*32*	*31*	*31*	*29*	**31**	*27*	*27*	*26*	*25*	**27**	*23*	*22*	*20*	*18*
V	**39**	*39*	*34*	*33*	*33*	*33*	*32*	*31*	*31*	*29*	**33**	*27*	*27*	*26*	*25*	*24*	*23*	*22*	*20*	*18*
M	**39**	*39*	*34*	*33*	*33*	*32*	*31*	*31*	*31*	*29*	**32**	*27*	*26*	*25*	*24*	*24*	*23*	*22*	*20*	*17*
I	**38**	*38*	*34*	*33*	*32*	**34**	*31*	*31*	*31*	*29*	*28*	*27*	*26*	*25*	*24*	*24*	*23*	*21*	*20*	*17*
N	**38**	*38*	*34*	*33*	*32*	*32*	*31*	*31*	*31*	*29*	*28*	*27*	*26*	*25*	*24*	*24*	*23*	*21*	*19*	*17*
S	**38**	*38*	*33*	*32*	*32*	*32*	*31*	*30*	*30*	*28*	*28*	*26*	*26*	*25*	*24*	*23*	*22*	*21*	*19*	*17*
H	**37**	*36*	*32*	*31*	*31*	*30*	*29*	*29*	*29*	*27*	*26*	*25*	*24*	*23*	*22*	*22*	*21*	*19*	*18*	*15*
T	**36**	*36*	*32*	*30*	*30*	*30*	*29*	*29*	*29*	*27*	*26*	*25*	*24*	*23*	*22*	*22*	*21*	*19*	*17*	*15*
C	**36**	*36*	*31*	*30*	*30*	*30*	*29*	*29*	*29*	*27*	*26*	*25*	*24*	*23*	*22*	*22*	*21*	*19*	*17*	*15*
Y	**34**	*34*	*30*	*29*	*28*	*28*	*27*	*27*	*27*	*25*	*24*	*23*	*22*	*21*	*20*	*20*	*19*	*17*	*15*	*13*
F	**34**	*33*	*29*	*28*	*28*	**24**	*26*	*26*	*26*	*24*	*23*	*22*	*21*	*20*	*19*	*19*	*18*	*16*	*15*	*12*
G	**33**	*33*	*29*	*27*	*27*	**26**	*26*	*26*	*26*	*24*	**27**	*22*	*21*	*20*	*19*	**25**	*18*	*16*	**20**	*12*
W	**32**	*32*	*27*	*26*	*26*	*26*	*25*	*24*	*24*	*22*	*21*	*20*	*20*	*19*	*18*	*17*	*16*	*15*	*13*	*11*

*The rows of amino acids are listed in order of their X position propensity for triple-helix formation, while the amino acids in columns are listed in order of their Y position propensity. The predicted T_m values are based on simple additivity of the stability of the residue in the X position and that in the Y position (see Persikov *et al.*, 2002).

Quantification of the intrinsic propensities and interactions for the most common sequences can now be used as a basis for predicting stability of peptides and local stability in collagens. Comparison of the predicted stability of more than 20 peptides with observed T_m values shows there is good agreement in the majority of the cases (Persikov *et al.*, 2002, 2005). This is an important asset in designing peptides that will form stable triple-helices, which are being used increasingly to map binding sites (see Section VII below) as well as for investigating interactions. It had been proposed some years ago that it would be possible to predict the local stability along collagen on the basis of individual tripeptide propensities (Bachinger and Davis, 1991), and these host-guest peptide data provide a realistic basis for such predictions.

VII. Ligand Binding

The binding of various ligands to the collagen triple helix is critical to biological function (Deprez *et al.*, 2000; Di Lullo *et al.*, 2002; Kadler, 1994; Knight *et al.*, 2000). Collagen binding to integrins and other cellular receptors mediates cell adhesion, while interaction of collagen with proteoglycans and other matrix molecules organizes the extracellular matrix, giving each tissue its distinctive mechanical properties. The binding of microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) to collagen mediates bacterial adhesion to the integrins of the host cell (Patti *et al.*, 1995). The turnover of collagen involves recognition of one specific site on the triple helix by matrix metalloproteinases (collagenase) and the biosynthesis of collagen involves binding of the chaperone Hsp47. Binding of ligands to collagen-like domains in host-defense proteins plays an important role in their physiological activity. The complement serine proteinases C1r and C1s bind to the collagenous domain of C1q during activation, while the MASP serine proteases bind to mannose binding lectin (MBL). Polyanionic ligands, including oxidized LDL, interact with the triple-helix domain of the macrophage scavenger receptor (MSR); the collagenous domain of asymmetric AChE binds to heparan sulfate in the basal lamina of neuromuscular junctions. As the list of proteins with collagen domains grows, so does the number of observed and proposed interactions. It is not clear how interactions mediated by a collagen triple helix compare to those found with globular proteins in terms of the nature of the interactions, degree of specificity, and strength of binding.

Rotary shadowing and transmission electron microscopy of collagen with bound ligand or antibodies have been used to roughly locate the binding site, and binding assays using cyanogen bromide fragments of

collagen have often narrowed down the region of interest (Di Lullo *et al*, 2002). The precise definition of the sequence involved has been determined using triple-helical peptides that cover the region of interest (Glattauer *et al.*, 1997; Knight *et al.*, 2000), or mutational analysis for smaller collagen-like domains that can be studied by recombinant DNA methodology (Doi *et al.*, 1993; Wallis *et al*, 2004). The known binding sites along type I collagen, both at a low and high resolution level, have been mapped by the group of San Antonio, in order to envision their relationships with each other and the D-periodic fibril organization (DiLullo *et al.*, 2002).

Although more than 50 molecules are listed as interacting with type I collagen (Di Lullo *et al.*, 2002) and ligands are known to bind to collagenous domains in host defense protein, it has proved difficult to define $(\text{Gly-X-Y})_n$ sequence binding sites within the triple helix. At this time, the specific sequence in the triple helix responsible for binding has been defined in a small number of cases: for integrin binding to type I collagen; the binding of a monoclonal antibody to type III collagen; heparin sulfate binding to the collagenous tail of acetylcholinesterase; ligand binding to MSR; Hsp47 to type I collagen; decorin binding to type I collagen; and MASP binding to MBL (Table V). In addition, the location of the unique

TABLE V
List of Known Binding Sequences in Collagens and Collagen-Like Proteins

Binding of	Binds to	Sequence	Reference
Type I collagen	Integrin $\alpha1\beta1$, $\alpha2\beta1$	GFOGER	Knight *et al.*, 2000
	HSP-47	Gly-X-Arg	Koide *et al.*, 2002
	Heparin	KGHRGF	Sweeney *et al.*, 1998
	Collagenase (cleavage site)	GLA or GIA	Fields, 1991
	Crosslinking site	Gly-X-Hyl-Gly-His-Arg-Gly	Kadler, 1994
Type III collagen	Monoclonal antibody	GLAGAOGLR	Glattauer *et al.*, 1997
Type IV collagen	Integrin $\alpha1\beta1$	$\alpha1$(IV)Asp461, $\alpha2$(IV)Arg461	Golbik *et al.*, 2000
Type V collagen	Heparin	GKPGPRGQRGPTGPRGER	Delacoux *et al.*, 2000
Collagenous tail of AchE	Heparin	GROGRKGRO, GROGKRGKQGQK	Deprez *et al.*, 2000
MBL collagen domain	MASP	GLRGLQGPOGKLGPOG	Wallis *et al.*, 2004

collagenase cleavage site and the two crosslinking sites have been well established, defining the collagen sequences that are recognized (Fields, 1991; Kadler, 1994).

In some cases there is evidence that a local decrease in stability or looseness promotes binding. The type III collagen epitope for a monoclonal antibody is flanked by destabilizing Gly-Gly-Y triplets (Shah *et al.*, 1997), and characterization of model peptides for the type IV cell adhesion has suggested conformational heterogeneity (Sacca *et al.*, 2003). Triple-helical peptide models of the collagenous tail of asymmetric acetylcholinesterase (AChE) show more effective binding to heparin when their stability is lower (Deprez *et al.*, 2000). The recognition sequence for serine proteinases (MASPs) on MBL appears to be just C-terminal to the destabilizing Gly-Gln-Gly interruption in the triple helix (Wallis *et al.*, 2004), and the collagenase cleavage site has a C-terminal locally unstable region (Fields, 1991).

A major advance in understanding collagen-ligand interactions came with the high-resolution structure of cocrystals in the I domain of $\alpha 2\beta 1$ integrin with a triple-helical peptide (Gly-Pro-Hyp)$_2$-Gly-Phe-Hyp-Gly-Glu-Arg-(Gly-Pro-Hyp)$_3$ containing the known type I collagen binding sequence (designated the IBP peptide) (Fig. 7; Emsley *et al.*, 2000). The

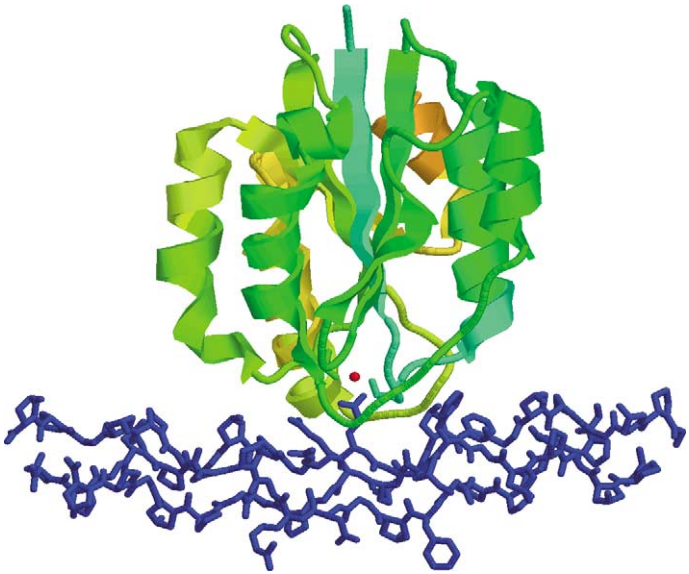


FIG. 7. Structure of the integrin I domain bound to the IBP triple-helix peptide.

three strands of this homotrimer—designated leading, middle, and trailing—have unique environments. Interactions with the I domain are largely mediated by the middle strand. The middle strand Glu completes the coordination sphere of a metal formed by the I domain MIDAS (metal ion-dependent adhesion site) motif, while the middle strand Arg forms a salt bridge to the I domain. The Phe residues of the middle and trailing strand make contact with the surface of the I domain. Peptide studies showed the Glu is essential and the Arg is required for high-affinity binding (Knight *et al.*, 2000). As discussed above in Section III.A, the IBP peptide has three distinct regions in terms of helix twist, and the kinking and bending at the junctions of these regions appears to play a critical role in optimizing interactions with the I domain (Emsley *et al*, 2004). The dominance of one strand in interactions with the integrin may have important implications for heterotrimeric collagen domains, and suggests the importance of recent strategies for synthesizing heterotrimeric collagen model peptides (e.g., see Sacca *et al.*, 2003).

VIII. Mutations and Disease

More than 1000 different mutations leading to human disorders have been observed in various types of collagens, and the clinical manifestations of these diseases reflect the specific tissue distribution of the particular collagen type that contains the mutation. Several excellent reviews of collagen diseases have been published (Byers and Cole, 2002; Myllyharju and Kivirikko, 2001, 2004). Well-characterized diseases include osteogenesis imperfecta (OI), where bone fragility results from mutations in type I collagen; Ehlers Danlos syndrome type IV, with aortic rupture as a result of mutations in type III collagen; Alport syndrome, with progressive renal failure resulting from basement membrane Type IV collagen mutations; and the dystrophic form of epidermolysis bullosa, with scarring and blistering of skin due to mutations in the skin anchoring fibrils Type VII collagen (Myllyharju and Kivirikko, 2001, 2004). Collagen mutations are generally dominant, because of their presence in multisubunit molecules and higher-order assemblies.

The first collagen genetic disease to be characterized was OI, a clinically heterogeneous disorder characterized by varying degrees of bone fragility (Byers and Cole, 2002). Cases are classified in four major categories, ranging from mild to perinatal lethal. OI results from mutations in the genes that code for the $\alpha 1$(I) or $\alpha 2$(I) chains of type I collagen, the major structural protein in bone. The majority of mutations involve single base substitutions in Gly codons in the triple helix, which lead to the replacement of a single Gly within the $(\text{Gly-X-Y})_{338}$ sequence by one of eight

bulkier amino acids (Ala, Arg, Asp, Cys, Glu, Ser, Trp, Val). More than 150 distinct mutations have been reported for different cases, and the mutation sites are located at varying sites along the length of both $\alpha 1$(I) and $\alpha 2$(I) collagen chains. The defective mineralization of bone collagen may be related to delayed collagen folding, a small decrease in helix stability, excess posttranslational modification, increased intracellular breakdown, reduced secretion, and abnormal fibril assembly (Byers and Cole, 2002). It is not known why different Gly substitution mutations lead to differing degrees of OI clinical severity, but it has been suggested to be related to the identity of the residue replacing Gly, the location of the mutation site with respect to the C-terminus, and the sequence immediately surrounding the mutation (Byers, 2001; Marini *et al.*, 1993). Disorders caused by single base substitutions in Gly codons within the triple helix are also common in other collagens, even in cases such as type IV collagen, where breaks in the Gly-X-Y repeating pattern are found normally (Hudson *et al.*, 2003).

Gly substitution mutations in the triple-helix domains of mannose-binding protein and C1q have also been associated with clinical disorders (Petry *et al.*, 1997; Turner, 2003). A Gly→Asp mutation in the fifth codon of the triple-helix domain of MBL is a common variant, reaching as high as 30% in some populations. This defect leads to MBL deficiency in the serum, resulting in susceptibility to infection, particularly in young children. Additional MBL mutations, Gly→Glu in the sixth codon of the triple helix and Arg→Cys in the Y position of the fourth codon are also observed. The Gly45Asp variant has been associated with defects in the oligomerization of MBL and also with binding to its associated MASPs, so that its function as host-defense is compromised, leading to increased infections (Wallis *et al.*, 2004).

Since Gly replacements are such a common feature in collagen diseases, it is important to understand the consequences of such a break in the Gly-X-Y repeating pattern for the collagen triple-helix structure, and to see if any molecular features correlate with clinical severity in OI. Peptides have been used to investigate the structural and energetic consequences of a Gly replacement on the collagen triple helix. A single replacement of a Gly by an Ala in $(\text{Pro-Hyp-Gly})_{10}$ was very destabilizing (Long *et al.*, 1993), and the crystal structure of this homotrimer G→A peptide showed only a small bulge at the Ala site, with a local disruption of NH...CO bonds, which are replaced by water-mediated hydrogen bonds (Bella *et al.*, 1994). The triple-helices in the Gly-Pro-Hyp sequences at both ends are well-ordered 7/2 helices, but the registration between the two ends is lost as a result of a slight untwisting at the replacement site. This loss of spatial coherence could lead to alterations in fibril formation and failure to

mineralize. More realistic peptide models, including a sequence of collagen with an OI site capped by C-terminal Gly-Pro-Hyp triplets, have been characterized by calorimetry, circular dichroism, and NMR spectroscopy (Baum and Brodsky, 1999). The introduction of a Gly to Ser or Ala replacement interrupts the C-to N-terminal folding of this peptide, suggesting that some renucleation event is needed to continue through Gly substitution sites.

The X-ray results on the Gly→Ala peptide have been complemented by computational studies to understand the effect of Gly replacements on the triple helix, largely carried out in Teri Klein's group (Klein and Huang, 1999; Mooney *et al.*, 2001; Radmer and Klein, 2004). Molecular dynamics and free energy calculations were done on the G→A peptide. All peptide models of mutations are limited by the use of homotrimers, while computational methods were used to evaluate the consequences of a more realistic model of having a Gly replacement in just one or two chains. More recently, molecular dynamic simulations were done on a sequence modeling the site of a lethal OI mutation, examining the effects of neighboring residues on the hydrogen bonding networks.

All Gly substitutions dramatically destabilize the triple helix in peptide models. Studies on host-guest peptides show the degree of destabilization depends on the identity of the residue replacing Gly. The residues, in order from least to most destabilizing, are: Ala < Ser < Cys < Arg <Val < Glu, Asp < Trp (Beck *et al.*, 2000). This order of destabilization correlates with the clinical severity of OI, demonstrated when two cases with different Gly replacements occur at the same site in α1(I) chains. For instance, Gly883Ser gives rise to a mild OI phenotype, while Gly883Asp gives rise to a lethal phenotype (Byers and Cole, 2002). The spectrum of amino acids replacing Gly in OI cases is significantly different from that predicted on the basis of nucleotide mutations rates and codon usage in type I collagen (Persikov *et al.*, 2004). This difference is most striking for the nonlethal cases, where the least and most destabilizing residues are underrepresented. This suggests that not every Gly replacement leads to a clinically detectable disorder, and supports the hypothesis that there is underrepresentation of the most destabilizing cases because they are not viable, and underrepresentation of the least destabilizing Gly→X substitutions because they are too mild to be clinically classified as OI. A schematic illustration of the number of different replacements observed in lethal and nonlethal cases of OI is shown in Fig. 8. The nonlethal and lethal categories represent observed data, which were scaled by the total number of Ser (the most frequent replacement in OI) and used to estimate the number that were undetectable for being too mild or too severe (Fig. 8).

more severe -------- less severe	less destabilizing ---------- more destabilizing						
Gly→	Ala	Ser	Cys	Arg	Val	Glu	Asp
Not observed Normal	[17] 65%	[0] 0%	[1] 5%	[1] 3%	[1] 4%	[0] 0%	[0] 0%
Observed Non-lethal	6 23%	30 51%	14 70%	15 41%	2 9%	1 9%	1 2%
Observed Lethal	3 12%	29 49%	5 25%	13 35%	10 44%	1 9%	8 18%
Not observed Not viable	[0] 0%	[0] 0%	[0] 0%	[8] 21%	[10] 43%	[9] 82%	[36] 80%

FIG. 8. Schematic showing the relationship between the degree of destabilization caused by a Gly replacement by a given residue and the degree of clinical severity seen in OI. The nonlethal and lethal phenotypes are observed data, while the normal and nonviable are inferred by scaling the total number of each possible replacement to the number of observed for Ser, as described in Persikov *et al.* (2004).

IX. CONCLUSIONS

The discovery of collagen triple-helix domains in a whole range of proteins, including bacteria and viruses, illustrates the adaptability of this motif to a range of biological functions. As more human disorders are found to be related to mutations in the collagen triple helix, the functional importance of some collagen types is clarified. These mutations provide a motivation for basic research on sequence-dependent alterations of triple-helix properties, and for investigating how such alterations may lead to a disease phenotype. This past decade has been marked by the determination of high-resolution structures of collagen triple-helix peptides and the availability of NMR data on specifically labeled residues. These studies have been complemented by dynamic, thermodynamic, and computational analyses. In the coming years, crystal structures of heterotrimers and of other ligands bound to collagen domains will be needed to further clarify the molecular basis of function.

X. ABBREVIATIONS AND NOTATION

The standard one-letter and three-letter amino acid notations are used, with hydroxyproline designated as Hyp in the three-letter code and O in the one-letter code. Fluoroproline is designated as Flp. Abbreviations used for different proteins include MBL, mannose binding lectin, also known as mannose binding protein; SP-A and SP-D, surfactant apoproteins A and D; MSR, macrophage scavenger receptor; IBP, integrin binding peptide; Scl1 and Scl2,

Streptococcus collagen-like protein 1 and 2; BclA, Bacillus collagen-like protein in anthrax. Peptides are indicated by the notation used in Table II, or by the single letter code for the residues other than Gly-Pro-Hyp triplets.

The individual collagen types are denoted by roman numerals, with individual chain types indicating alpha chains. For example, $\alpha 1$(V) indicates the $\alpha 1$ chain of type V collagen.

The symmetry of the triple helix is designated here as 10/3 (twist angle of 36 degrees, 10-fold symmetry, which is also designated as 10_7 in crystallographic screw symmetry notation) and 7/2 (twist angle of 51.4 degrees, sevenfold symmetry, which is also designated as 7_5 in crystallographic screw symmetry notation).

Acknowledgments

This work is supported by National Institutes of Health (NIH) grant GM60048 (to B.B.), by the Michael Geisman OI Research Fellowship (to A.P.), and a Ruth Kirschstein NRSA NIH award (to B.B.).

References

Bachinger, H. P., and Davis, J. M. (1991). Sequence specific thermal stability of the collagen triple-helix. *Int. J. Biol. Macromol.* **13,** 152–156.

Bann, J. G., and Bachinger, H. P. (2000). Glycosylation/hydroxylation-induced stabilization of the collagen triple helix. 4-trans-hydroxyproline in the Xaa position can stabilize the triple helix. *J. Biol. Chem.* **275,** 24466–24469.

Batchelder, L. S., Sullivan, C. E., Jelinski, L. W., and Torchia, D. A. (1982). Characterization of leucine side-chain reorientation in collagen fibrils by solid state 2H NMR. *Proc. Natl. Acad. Sci.* **79,** 386–389.

Baum, J., and Brodsky, B. (1999). Folding of peptide models of collagen and misfolding in disease. *Curr. Opin. Struct. Biol.* **9,** 122–128.

Beck, K., Chan, V. C., Shenoy, N., Kirkpatrick, A., Ramshaw, J. A., and Brodsky, B. (2000). Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine. *Proc. Natl. Acad. Sci.* **97,** 4273–4278.

Bella, J., and Berman, H. M. (1996). Crystallographic evidence for C α-H$\cdots$O=C hydrogen bonds in a collagen triple helix. *J. Mol. Biol.* **264,** 734–742.

Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994). Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* **266,** 75–81.

Bella, J., Brodsky, B., and Berman, H. M. (1995). Hydration structure of a collagen peptide. *Structure* **3,** 893–906.

Berendsen, H. J. C., and Migchelsen, C. (1965). Hydration structure of fibrous macromolecules. *Ann. N.Y. Acad. Sci.* **125,** 365–379.

Berisio, R., Vitagliano, L., Mazzarella, L., and Zagari, A. (2001). Crystal structure of a collagen-like polypeptide with repeating sequence Pro-Hyp-Gly at 1.4Å resolution: Implications for collagen hydration. *Biopolymers* **56,** 8–13.

Berisio, R., Vitagliano, L., Mazzarella, L., and Zagari, A. (2002). Recent progress on collagen triple helix structure, stability and assembly. *Prot. Pept. Lett.* **9,** 107–116.

Berisio, R., Granata, V., Vitagliano, L., and Zagari, A. (2004). Imino acids and collagen triple helix stability: Characterization of collagen-like peptides containing Hyp-Hyp-Gly sequence repeats. *J. Am. Chem. Soc.* **126,** 11402–11403.

Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L., and Raines, R. T. (2001). Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc.* **123,** 777–778.

Burjanadze, T. V. (2000). New analysis of the phylogenetic change of collagen thermostability. *Biopolymers* **53,** 523–528.

Byers, P. H. (2001). Folding defects in fibrillar collagens. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356,** 151–158.

Byers, P. H., and Cole, W. G. (2002). Osteogenesis Imperfecta. *In* ''Connective tissue and its hereditable disorders,'' (P. M. Royce and B. Steinmann, Eds.), pp. 385–430. Wiley-Liss., New York.

Chung, L., Dinakarpandian, D., Yoshida, N., Lauer-Fields, J. L., Fields, G. B., Visse, R., and Nagase, H. (2004). Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J.* **23,** 3020–3030.

Cowan, P. M., McGavin, S., and North, A. C. T. (1955). The polypeptide configuration of collagen. *Nature* **176,** 1062–1064.

De Rider, M. L., Wilkens, S. J., Waddell, M. J., Bretcher, L. E., Weinhold, F., Raines, R. T., and Markely, J. L. (2002). Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. *J. Am. Chem. Soc.* **124,** 2497–2505.

Delacoux, F., Fichard, A., Cogne, S., Garrone, R., and Ruggiero, F. (2000). Unraveling the amino acid sequence crucial for heparin binding to collagen V. *J. Biol. Chem.* **275,** 29377–29382.

Deprez, P., Doss-Pepe, E., Brodsky, B., and Inestrosa, N. C. (2000). Interaction of the collagen-like tail of asymmetric acetylcholinesterase with heparin depends on triple-helical conformation, sequence and stability. *Biochem. J.* **350,** 283–290.

Di Lullo, G. A., Sweeney, S. M., Korkko, J., Ala-Kokko, L., and San Antonio, J. D. (2002). Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *J. Biol. Chem.* **277,** 4223–4231.

Doi, T., Higashino, K., Kurihara, Y., Wada, Y., Miyazaki, T., Nakamura, H., Uesugi, S., Imanishi, T., Kawabe, Y., Itakura, H., Yazaki, Y., Matsumoto, A., and Kodama, T. (1993). Charged collagen structure mediates the recognition of negatively charged macromolecules by macrophage scavenger receptors. *J. Biol. Chem.* **268,** 2126–2133.

Doi, M., Nishi, Y., Uchiyama, S., Nishiuchi, Y., Nakazawa, T., Ohkubo, T., and Kobayashi, Y. (2003). Characterization of collagen model peptides containing 4-fluoroproline; $(4(S)\text{-fluoroproline-Pro-Gly})_{10}$ forms a triple helix, but $(4(R)\text{-fluoroproline-Pro-Gly})_{10}$ does not. *J. Am. Chem. Soc.* **125,** 9922–9923.

Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000). Structural basis of collagen recognition by integrin $\alpha 2\beta 1$. *Cell* **101,** 47–56.

Emsley, J., Knight, C. G., Farndale, R. W., and Barnes, M. J. (2004). Structure of the integrin alpha2beta1-binding collagen peptide. *J. Mol. Biol.* **335,** 1019–1028.

Engel, J., and Prockop, D. J. (1998). Does bound water contribute to the stability of collagen? *Matrix Biol.* **17,** 679–680.

Exposito, J. Y., Cluzel, C., Garrone, R., and Lethias, C. (2002). Evolution of collagens. *Anat. Rec.* **268,** 302–316.

Fan, P., Li, M., Brodsky, B., and Baum, J. (1993). Backbone dynamics of (Pro-Hyp-Gly)10 and a designed collagen-like triple-helical peptide by 15N NMR relaxation and hydrogen-exchange measurements. *Biochemistry* **32,** 13299–13309.

Fields, GB. (1991). A model for interstitial collagen catabolism by mammalian collagenases. *J. Theor. Biol.* **153,** 585–602.

Fields, G. B., and Prockop, D. J. (1996). Perspectives on the synthesis and application of triple-helical, collagen-model peptides. *Biopolymers* **40,** 345–357.

Fraser, R. D. B., and MacRae, T. P. (1973). ''Conformation of Fibrous Proteins.'' Academic Press, New York.

Fraser, R. D. B., MacRae, T. P., and Suzuki, E. (1979). Chain conformation in the collagen molecule. *J. Mol. Biol.* **129,** 463–481.

Fraser, R. D. B., MacRae, T. P., Miller, A., and Suzuki, E. (1983). Molecular conformation and packing in collagen fibrils. *J. Mol. Biol.* **167,** 497–521.

Fraser, R. D. B., MacRae, T. P., Chew, M. W. K, and Squire, J. M. (1987). Collagen and Elastin. *In* ''Fibrous Protein Structure,'' (J. M. Squire and P. J. Vibert, Eds.), pp. 173–191. Academic Press, London.

Glattauer, V., Werkmeister, J. A., Kirkpatrick, A., and Ramshaw, J. A. M. (1997). Identification of the epitope for a monoclonal antibody that blocks platelet aggregation by type III collagen. *Biochem. J.* **323,** 45–49.

Golbik, R., Eble, J. A., Ries, A., and Kuhn, K. (2000). The spatial orientation of the essential amino acid residues arginine and aspartate within the alpha1beta1 integrin recognition site of collagen IV has been resolved using fluorescence resonance energy transfer. *J. Mol. Biol.* **297,** 501–509.

Goodman, M., Bhumralkar, M., Jeffersohn, E. A., Kwak, J., and Locardi, E. (1998). Collagen mimetics. *Biopolymers* **47,** 127–142.

Grigera, J. R., and Berendsen, H. J. C. (1979). The molecular details of collagen hydration. *Biopolymers* **18,** 47–57.

Gustavson, K. H. (1955). The function of hydroxyproline in collagen. *Nature* **175,** 70–74.

Harrington, W. F., and von Hippel, P. F. (1961). The Structure of Collagen and Gelatin. *In* ''Advances in Protein Chemistry,'' (C. B. Anfinsen, Ed.), Vol. 16, pp. 1–138. Academic Press, New York.

Heidemann, E., and Roth, W. (1982). Synthesis and investigation of collagen model peptides. *Adv. Polym. Sci.* **43,** 142–203.

Hodges, J. A., and Raines, R. T. (2003). Stereoelectronic effects on collagen stability: The dichotomy of 4-fluoroproline diastereomers. *J. Am. Chem. Soc.* **125,** 9262–9263.

Holmgren, S. K., Bretscher, L. E., Taylor, K. M., and Raines, R. T. (1999). A hyperstable collagen mimic. *Chem. Biol.* **6,** 63–70.

Hudson, B. G., Tryggvason, K., Sundarmoorthy, M., and Neilson, E. G. (2003). Alport's syndrome, Goodpasture's syndrome, and type IV collagen. *N. Engl. J. Med.* **348,** 2543–2556.

Inouye, K., Sakakibara, S., and Prockop, D. J. (1976). Effects of the stereo-configuration of the hydroxyl group in 4-hydroxyproline on the triple-helical structures formed by homogenous peptides resembling collagen. *Biochim. Biophys. Acta* **420,** 133–141.

Inouye, K., Kobayashi, Y., Kyogoku, Y., Kishida, Y., Sakakibara, S., and Prockop, D. J. (1982). Synthesis and physical properties of (Hydroxyproline-Proline-Glycine)$_{10}$: Hydroxyproline in the X-position decreases the melting temperature of the collagen triple helix. *Arch. Biochem. Biophys.* **219,** 198–203.

Jenkins, C. L., and Raines, R. T. (2002). Insights on the conformational stability of collagen. *Nat. Prod. Rep.* **19,** 49–59.

Jenkins, C. L., Bretscher, L. E., Guzei, I. A., and Raines, R. T. (2003). Effect of 3-hydroxyproline residues on collagen stability. *J. Am. Chem. Soc.* **125,** 6422–6427.

Kadler, K. (1994). Extracellular matrix fibril forming collagen. *Protein Profile* **1,** 519–638.

Kassner, A., Tiedemann, K., Notbohm, H., Ludwig, T., Morgelin, M., Reinhardt, D. P., Chu, M.-L., Bruckner, P., and Grassel, S. (2004). Molecular structure and interaction of recombinant human type XVI collagen. *J. Mol. Biol.* **339,** 835–853.

Kawashima, H., Watanabe, N., Hirose, M., Sun, X., Atarashi, K., Kimura, T., Shikata, K., Matsuda, M., Ogawa, D., Heljasvaara, R., Rehn, M., Pihlajaniemi, T., and Miyasaka, M. (2003). Collagen XVIII, a basement membrane heparan sulfate proteoglycan, interacts with L-selectin and monocyte chemoattractant protein-1. *J. Biol. Chem.* **278,** 13069–13076.

Kielty, C. M., and Grant, M. E. (2002). The Collagen Family: Structure Assembly and Organization in the Extracellular Matrix. *In* "Connective Tissue and its Heritable Disorders," (P. M. Royce and B. Steinmann, Eds.), pp. 159–222. Wiley-Liss, New York.

Kilchherr, E., Hofmann, H., Steigemann, W., and Engel, J. (1985). Structural model of the collagen-like region of C1q comprising the kink region and the fibre-like packing of the six triple helices. *J. Mol. Biol.* **186,** 403–415.

Klein, T. E., and Huang, C. C. (1999). Computational investigations of structural changes resulting from point mutations in a collagen-like peptide. *Biopolymers* **49,** 167–183.

Knight, C. G., Morton, L. F., Peachey, A. R., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (2000). The collagen-binding A-domains of integrins $\alpha1\beta1$ and $\alpha2\beta1$ recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens. *J. Biol. Chem.* **275,** 35–40.

Koide, T., Takahara, Y., Asada, S., and Nagata, K. (2002). Xaa-Arg-Gly triplets in the collagen triple helix are dominant binding sites for the molecular chaperone HSP47. *J. Biol. Chem.* **277,** 6178–6182.

Kramer, R. Z., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (1999). Sequence dependent conformational variations of collagen triple-helical structure. *Nat. Struct. Biol.* **6,** 454–457.

Kramer, R. Z., Venugopal, M. G., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (2000). Staggered molecular packing in crystals of a collagen-like peptide with a single charged pair. *J. Mol. Biol.* **301,** 1191–1205.

Kramer, R. Z., Bella, J., Brodsky, B., and Berman, H. M. (2001). The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence. *J. Mol. Biol.* **311,** 131–147.

Krimm, S., and Kuroiwa, K. (1968). Low temperature infrared spectra of polyglycines and CH. . .O═C hydrogen bonding in polyglycine II. *Biopolymers* **6,** 401–407.

Leikin, S., Rau, D. C., and Parsegian, V. A. (1994). Direct measurement of forces between self-assembled proteins. Temperature-dependent exponential forces between collagen triple-helices. *Proc. Natl. Acad. Sci.* **91,** 276–280.

Leikin, S. Rau, D. C., and Parsegian, V. A. (1995). Temperature-favoured assembly of collagen is driven by hydrophilic not hydrophobic interactions. *Nat. Struct. Biol.* **2,** 205–210.

Leikin, S., Parsegian, V. A., Yang, W.-H., and Walrafen, G. E. (1997). Raman spectral evidence for hydration forces between collagen triple-helices. *Proc. Natl. Acad. Sci.* **94,** 11312–11317.

Liu, J. (2000). X-ray crystallographic studies of two nucleic-acid binding protein/complex systems and one collagen-like peptide. *Ph.D. thesis, Rutgers University.*

Long, C. G., Braswell, E., Zhu, D., Apigo, J., Baum, J., and Brodsky, B. (1993). Characterization of collagen-like peptides containing interruptions in the repeating GLy-X-Y sequence. *Biochemistry* **32,** 11688–11695.

Lu, J., the, C., Kishore, U., and Reid, K. B. M. (2002). Collectins and ficolins: Sugar pattern recognition molecules of the mammalian innate immune system. *Biochim. Biophys. Acta* **1572,** 387–400.

Marini, J. C., Lewis, M. B., Wang, Q., Chen, K. J., and Orrison, B. M. (1993). Serine for glycine substitutions in type I collagen in two cases of type IV osteogenesis imperfecta (OI). Additional evidence for a regional model of OI pathophysiology. *J. Biol. Chem.* **268,** 2667–2673.

Matsuzawa, Y., Funahashi, T., Kihara, S., and Shimomura, I. (2004). Adiponectin and metabolic syndrome. *Arterioscler. Thromb. Vasc. Biol.* **24,** 29–33.

McAlinden, A., Smith, T. A., Sandell, L. J., Ficheux, D., Parry, D. A. D., and Hulmes, D. J. S. (2003). Alpha-helical coiled coil oligomerization domains are almost ubiquitous in the collagen superfamily. *J. Biol. Chem.* **278,** 42200–42207.

Melacini, G., Bonvin, A. M., Goodman, M., Boelens, R., and Kaptein, R. (2000). Hydration dynamics of the collagen triple helix by NMR. *J. Mol. Biol.* **300,** 1041–1049.

Mizuno, K., Hayashi, T., Peyton, D. H., and Bachinger, H. P. (2004). The peptides acetyl-(Gly-3(S)Hyp-4(R)Hyp)$_{10}$-NH2 and acetyl-(Gly-Pro-3(S)Hyp)$_{10}$-NH2 do not form a collagen triple helix. *J. Biol. Chem.* **279,** 282–287.

Mooney, S. D., Huang, C. C., Kollman, P. A., and Klein, T. E. (2001). Computed free energy differences between point mutations in a collagen-like peptide. *Biopolymers* **58,** 347–353.

Myllyharju, J., and Kivirikko, K. I. (2001). Collagen and collagen-related diseases. *Ann. Med.* **33,** 7–21.

Myllyharju, J., and Kivirikko, K. I. (2004). Collagens, modifying enzymes, and their mutations in humans, flies and worms. *Trends Genet.* **20,** 33–43.

Okuyama, K., Arnott, S., Takayanagi, M., and Kakudo, M. (1981). Crystal and molecular structure of a collagen-like polypeptide (Pro-Pro-Gly)$_{10}$. *J. Mol. Biol.* **152,** 427–443.

Patti, J. M., House-Pompeo, K., Boles, J. O., Garza, N., Gurusiddappa, S., and Hook, M. (1995). Critical residues in the ligand-binding site of the Staphylococcus aureus collagen-binding adhesion (MSCRAMM). *J. Biol. Chem.* **270,** 12005–12011.

Perret, S., Merle, C., Bernocco, S., Berland, P., Garrone, R., Hulmes, D. J., Theisen, M., and Ruggiero, F. (2001). Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation. *J. Biol. Chem.* **276,** 43693–43698.

Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A., and Brodsky, B. (2000). Amino acid propensities for the collagen triple-helix. *Biochemistry* **39,** 14960–14967.

Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A., and Brodsky, B. (2002). Peptide investigations of pairwise interactions in the collagen triple-helix. *J. Mol. Biol.* **316,** 385–394.

Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2005). Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability. *Biochemistry* **44,** 1414–1422.

Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A., and Brodsky, B. (2003). Triple-helix propensity of hydroxyproline and fluoroproline: Comparison of host-guest and repeating tripeptide collagen models. *J. Am. Chem. Soc.* **125,** 11500–11501.

Persikov, A. V., Pillitteri, R. J., Amin, P., Schwarze, U., Byers, P. H., and Brodsky, B. (2004). Stability related bias in residues replacing glycines within the collagen triple helix (Gly-Xaa-Yaa) in inherited connective tissue disorders. *Hum. Mut.* **24,** 330–337.

Petry, F., Hauptmann, G., Goetz, J., Grosshans, E., and Loos, M. (1997). Molecular basis of a new type of C1q deficiency associated with a non-functional low molecular weight (LMW) C1q: Parallels and differences to other known genetic C1q defects. *Immunopharmacology* **38,** 189–201.

Privalov, P. L. (1982). Stability of proteins: Proteins which do not present a single cooperative system. *Adv. Prot. Chem.* **35,** 1–104.

Radmer, R. J., and Klein, T. E. (2004). Severity of osteogenesis imperfecta and structure of a collagen-like peptide modeling a lethal mutation site. *Biochemistry* **43,** 5314–5323.

Ramachandran, G. N. (1967). Structure of Collagen at the Molecular Level. *In* ''Treatise on Collagen,'' (G. N. Ramachandran, Ed.), pp. 103–183. Academic Press, New York.

Ramachandran, G. N., and Kartha, G. (1954). Structure of collagen. *Nature* **174,** 269–270.

Ramachandran, G. N., and Kartha, G. (1955). Structure of collagen. *Nature* **176,** 593–595.

Ramachandran, G. N., and Chandrasekharan, R. (1968). Interchain hydrogen bonds via bound water molecules in the collagen triple-helix. *Biopolymers* **6,** 1649–1658.

Ramachandran, G. N., Bansal, M., and Bhatnagar, R. S. (1973). A hypothesis on the role of hydroxyproline in stabilizing collagen structure. *Biochim. Biophys. Acta* **322,** 166–171.

Rasmussen, M., Jacobsson, M., and Bjorck, L. (2003). Genome-based identification and analysis of collagen-related structural motifs in bacterial and viral proteins. *J. Biol. Chem.* **278,** 32313–32316.

Rich, A., and Crick, F. H. C. (1955). The structure of collagen. *Nature* **176,** 915–916.

Rich, A., and Crick, F. H. C. (1961). The molecular structure of collagen. *J. Mol. Biol.* **3,** 483–506.

Rigby, B. J., and Robinson, M. S. (1975). Thermal transitions in collagen and the preferred temperature range of animals. *Nature* **253,** 277–279.

Rosenbloom, J., Harsh, M., and Jimenez, S. (1973). Hydroxyproline content determines the denaturation temperature of chick tendon collagen. *Arch. Biochem. Biophys.* **158,** 478–484.

Sacca, B., Fiori, S., and Moroder, L. (2003). Studies of the local conformational properties of the cell-adhesion domain of collagen type IV in synthetic heterotrimeric peptides. *Biochemistry* **42,** 3429–3436.

Sakikabara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y., and Prockop, D. J. (1973). Synthesis of $(\text{Pro-Hyp-Gly})_n$ of defined molecular weights. Evidence for stabilization of collagen triple-helix by hydroxyproline. *Biochim. Biphys. Acta* **303,** 198–202.

Sarkar, S. K., Sullivan, C. E., and Torchia, D. A. (1983). Solid state 13C NMR study of collagen molecular dynamics in hard and soft tissues. *J. Biol. Chem.* **258,** 9762–9767.

Sarkar, S. K., Hiyama, Y., Niu, C. H., Young, P. E., Gerig, J. T., and Torchia, D. A. (1987). Molecular dynamics of collagen side chains in hard and soft tissues. A multinuclear magnetic resonance study. *Biochemistry* **26,** 6793–6800.

Shah, N. K., Sharma, M., Kirkpatrick, A., Ramshaw, J. A. M., and Brodsky, B. (1997). Gly-Gly containing triplets of low stability adjacent to a type III collagen epitope. *Biochemistry* **36,** 5878–5883.

Shirai, H., Murakami, T., Yamada, Y., Doi, T., Hamakubo, T., and Kodama, T. (1999). Structure and function of type I and II macrophage scavenger receptors. *Mech. Ageing Dev.* **111,** 107–121.

Siebold, B., Qian, R. A., Glanville, R. W., Hofmann, H., Deutzmann, R., and Kuhn, K. (1987). Identification of the cleavage sites by a hemorrhagic metalloproteinase in type IV collagen. Construction of a model for the aggregation and cross-linking region (7S domain) of type IV collagen based upon an evaluation of the primary structure of the alpha 1 and alpha 2 chains in this region. *Eur. J. Biochem.* **168,** 569–575.

Simon-Lukasik, K. V., Persikov, A. V., Brodsky, B., Ramshaw, J. A., Laws, W. R., Ross, A., and Ludescher, R. D. (2003). Fluorescence determination of tryptophan side-chain accessibility and dynamics in triple-helical collagen-like peptides. *Biophys. J.* **84,** 501–508.

Steplewski, A., Majsterek, I., McAdams, E., Rucker, E., Brittingham, R. J., Ito, H., Kazuya, H., Adachi, E., Jimenez, S. A., and Fertala, A. (2004). Thermostability gradient in the collagen triple-helix reveals its multi-domain structure. *J. Mol. Biol.* **338,** 989–998.

Stetefeld, J., Frank, S., Jenny, M., Schulthess, T., Kammerer, R. A., Boudko, S., Landwehr, R., Okuyama, K., and Engel, J. (2003). Collagen stabilization at atomic level: Crystal structure of designed $(GlyProPro)_{10}$-foldon. *Structure* **11,** 339–346.

Stultz, C. M. (2002). Localized unfolding of collagen explains collagenase cleavage near imino-poor sites. *J. Mol. Biol.* **319,** 997–1003.

Suzuki, E., Fraser, R. D. B., and MacRae, T. P. (1980). Role of hydroxyproline in the stabilization of the collagen molecule via water molecules. *Int. J. Biol. Macrom.* **2,** 54–56.

Sweeney, S. M., Guy, C. A., Fields, G. B., and San Antonio, J. D. (1998). Defining the domains of type I collagen involved in heparin-binding and endothelial tube formation. *Proc. Natl. Acad. Sci.* **95,** 7275–7280.

Sylvestre, P., Couture-Tose, E., and Mock, M. (2003). Polymorphism in the collagen-like region of the Bacillus anthracis BclA protein leads to variation in exosporium filament length. *J. Bacteriol.* **185,** 1555–1563.

Traub, W., and Piez, K. A. (1971). The chemistry and structure of collagen. *Adv. Prot. Chem.* **25,** 243–352.

Turner, M. W. (2003). The role of mannose binding lectin in health and disease. *Mol. Immunol.* **40,** 423–429.

Vitagliano, L., Nemethy, G., Zagari, A., and Scheraga, H. A. (1993). Stabilization of the triple-helical structure of natural collagen by side chain interactions. *Biochemistry* **32,** 7354–7359.

Vitagliano, L., Berisio, R., Mazzarella, L., and Zagari, A. (2001). Structural bases of collagen stabilization induced by proline hydroxylation. *Biopolymers* **58,** 459–464.

Wallis, R., Shaw, J. M., Uitdehaag, J., Chen, C. B., Torgersen, D., and Drickamer, K. (2004). Location of the serine protease-binding sites in the collagen-like domain of mannose-binding protein. *J. Biol. Chem.* **279,** 14065–14073.

Xu, Y., Keene, D. R., Bujnicki, J. M., Hook, M., and Lukomski, S. (2002). Streptococcal Scl1 and Scl2 proteins form collagen-like triple-helices. *J. Biol. Chem.* **277,** 27312–27318.

Xu, Y., Hyde, T., Wang, X., Bhate, M., and Brodsky, B. (2003). NMR and CD spectroscopy show that imino acid restriction of the unfolded state leads to efficient folding. *Biochemistry* **42,** 8696–8703.

Yonath, A., and Traub, W. (1975). Crystallization of cyanogen bromide peptides from chick cartilage collagen. *FEBS Lett.* **57,** 93–95.

Yee, R. W., Englander, S. W., and Von Hippel, P. H. (1974). Native collagen has a two-bonded structure. *J. Mol. Biol.* **83,** 1–16.

COLLAGEN FIBRIL FORM AND FUNCTION

By T. J. WESS

Structural Biophysics Division, School of Optometry and Vision Science, Cardiff University, Cardiff, Wales, United Kingdom

I. Introduction 342
II. The Collagens that Constitute Fibrils 342
A. Type I Collagen-Rich Fibrillar Structures 343
B. Type II Collagen-Rich Fibrillar Structures 344
III. Fibrillar Molecular Packing 345
IV. Lateral Packing in Fibrillar Collagens 350
V. Novel Fibrillar Structures 354
VI. Subfibrillar Structures 356
VII. Fibril Surface 356
VIII. Factors that May Regulate Fibril Growth and Size 357
IX. Distribution of Fibril Sizes 359
X. Suprafibrillar Architectures 361
XI. Mechanical Properties of Collagen-Rich Tissues 365
XII. Conclusions 366
References 367

Abstract

The majority of collagen in the extracellular matrix is found in a fibrillar form, with long slender filaments each displaying a characteristic ~67 nm D-repeat. Here they provide the stiff resilient part of many tissues, where the inherent strength of the collagen triple helix is translated through a number of hierarchical levels to endow that tissue with its specific mechanical properties. A number of collagen types have important structural roles, either comprising the core of the fibril or decorating the fibril surface to give enhanced functionality. The architecture of subfibrillar and suprafibrillar structures (such as microfibrils), lateral crystalline and liquid crystal ordering, interfibrillar interactions, and fibril bundles is described. The fibril surface is recognized as an area that contains a number of intimate interactions between different collagen types and other molecular species, especially the proteoglycans. The interplay between molecular forms at the fibril surface is discussed in terms of their contribution to the regulation of fibril diameter and their role in interfibrillar interactions.

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70010-8

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

I. Introduction

Collagen constitutes the major proteinaceous weight of the mammalian body, where it is commonly found as long, slender fibrillar structures that are most easily recognized by a ~67 nm periodicity. The collagen fibril serves as one of the prominent scaffolding structures utilized in animals, where the strength of an individual ropelike collagen molecule relates directly to the structural integrity and strength of the tissue. Collagen fibrils are substantial constituents of skin, tendon, bone, ligament, cornea, and cartilage, where the fundamental tensile properties of the fibril are finely tuned to serve bespoke biomechanical, structural, and mechanotransductory signaling roles. Many of these properties derive from the structural organization within a fibril, where the organization and topology of the collagen molecules ensure strong intermolecular interactions.

The presence of subfibrillar organization may point to structural levels of organization that are required for the successful mechanical response of fibrillar collagens, and may also be part of the inevitable balance between crystallinity and disorder within a biological polymer. The presence of different collagen types within a single fibril is a structural prerequisite in many tissues. However, the necessity for heterotypic fibrillar structures may point to fine tuning of the structural properties in a composite, such as fibril size regulation, dispersion of crystallinity, and interfibrillar communication. These structure-function relationships are still being resolved.

The overall properties and morphology of a fibril are as important as its internal organization. For example, the surface of a fibril is a complex area that contains collagen molecules and a variety of proteoglycans. These dictate the interaction between fibrils and specify the environment of partner macromolecules. They are also important in restricting fibril growth and permitting fusion to occur. The diameter and overall slender tapering of collagen fibrils has significance in determining the macroscopic mechanical properties of the tissues. The basis of fibrils existing either as a fused mass in a tissue or as discontinuous discrete structures is still the subject of debate.

The purpose of this Chapter is to review the major features of collagen fibrils and to demonstrate the current state of understanding the way collagen fibrils produce functional materials.

II. The Collagens that Constitute Fibrils

Collagen fibrils contain a number of collagen gene products. These can be conveniently divided into two classes: (a) fibril-forming collagens; and (b) fibril-associated collagens with interrupted triple helices (FACIT

collagens). In terms of incorporation into fibril structure, the fibril-forming collagens (Types I, II, III, V, and XI) usually constitute the majority of molecules. Their structure is dominated by the 300 nm triple-helical structure, though some differences in the size and complexity of the telopeptides (the ends of the molecule) do occur with collagen type. FACIT collagens are a far more heterogeneous family of molecules (e.g., Types IX, XII, XIV, XVI and XIX). Their additional structural complexity is thought to be due to the requirements for interactions between each fibril and the cell/matrix environment. To that end, FACITs usually occupy the surface of fibrils and some have a transient interaction with fibrils during development.

A further subdivision of the fibril-forming collagens can be made by examining the patterns of occurrence of such collagens in fibrils. Types I and II collagens form two classes of fibrils, in which they are the major form of collagen present. For the purpose of this review, such fibrils will be named Type-I-rich fibrils and Type-II-rich fibrils, respectively. Fibril-forming collagen Types III and V are preferentially associated with Type-I-rich fibrils, whereas Type-II-rich fibrils frequently contain Types IX and XI collagen. Fibrils comprising Types II and III collagen have also been observed in cartilage (Young *et al.*, 2000a). Many fibrillar structures are therefore described as heterotypic. Fibrillar composition can now be detected by very sensitive techniques, such as infrared matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (IR-MALDI-TOF-MS) (Dreisewerd *et al.*, 2004).

A. *Type I Collagen-Rich Fibrillar Structures*

Fibrillar structures that consist primarily of Type I collagen, such as those from rat tail tendon, appear to have distinct structural features such as a wide distribution of fibril diameters and a greater degree of internal crystallinity. It is also of importance that Type I collagen is a heterotrimeric structure, where the presence of the $\alpha 2$ chain is critical to proper fibril formation (McBride *et al.*, 1997). Type I-rich fibrils in tissues such as skin, however, are heterotypic and contain significant amounts of Type III collagen, typically around 20% by mass. The length of the triple helix in Type III collagen is slightly longer than in Type I although, in contrast, the telopeptides are slightly shorter. The presence of Type III collagen is often associated with tissues with regulated fibril diameters. Although Type III collagen can be buried within the fibril, slowly processed N-terminal propeptides may persist on its surface. Crosslinks between Type I and Type III collagen have also been identified, indicating specific interactions within the fibril (Henkel and Glanville, 1982).

Type V collagen is coassembled with Type I collagen in fibrils such as those in cornea, and is thought to be one of the factors responsible for the small, uniform fibrillar diameter (~25 nm) characteristic of this tissue (Linsenmayer *et al.*, 1993). The entire triple-helical domain of the Type V collagen molecules is believed to be buried within the fibril. The Type I collagen molecules are thus present at the fibril surface (Chanut-Delalande *et al.*, 2001), along with the retained N-terminal domains of the Type V collagen. The latter are believed to extend outward through the gap zones. A significant feature of the triple-helical domain of Type V collagen is the high content of glycosylated hydroxylysine residues (10 times higher than for Type I collagen); this must have a significant influence on both the intermolecular and interfibrillar interactions of the triple-helical domain of the Type V collagen molecule (Mizuno *et al.*, 2001).

Type XIV collagen contains multiple domains and is capable of interacting with collagen fibrils and other extracellular matrix components. Immunoelectron microscopy has shown that Type XIV collagen is fibril-associated with a periodicity of 67 nm, indicating specific interactions that may alter during tissue development. Changes in splice variant expression of the gene product suggest that different functional forms of Type XIV collagen can be present. This allows modified interactions with fibrils during development that facilitate the transition from growing fibril intermediates to mature fibrils (Young *et al.*, 2000b).

In the corneal stroma, Type XII collagen may be organized along the collagen fibrils in a uniform head-to-tail pattern (Wessel *et al.*, 1997). From its similarity to Types XII and XIV collagen, Type XX is also expected to bind to collagen fibrils (Koch *et al.*, 2001).

B. Type II Collagen-Rich Fibrillar Structures

Type II collagen forms the main scaffold for the second class of fibrils described here. Far less is known about the molecular structure within Type II-rich than Type I-rich fibrils. This results from the composition heterogeneity in tissues rich in Type II collagen. The fibrils usually also lack crystallinity, which has limited the interpretation of X-ray fiber diagrams. Also, fewer electron micrograph studies have been conducted (Ronziere *et al.*, 1987, 1998). Type II collagen and its associated minor collagens tend to be glycosylated to a greater extent than Type I-rich fibrils, which may provide part of the basis for the structural differences that occur. Cartilage fibrils contain Type II collagen as a major constituent, but the presence of additional components, minor collagens, and

noncollagenous glycoproteins is thought to be crucial for modulating several fibril properties.

Collagen XI, a heterotrimeric molecule, is found predominantly in heterotypic cartilage fibrils, where it is incorporated into Type II-rich fibrils and is involved in the regulation of fibrillogenesis. The partly processed nonhelical N-terminal regions are important in altering fibril surface properties (Blaschke *et al.*, 2000). Coassembly of collagens I and XI has also been found in fibrils of several normal and pathologically altered tissues, including fibrous cartilage, bone, and osteoarthritic joints (Hansen and Bruckner, 2003). The fibril structures appear to be significantly different from fibrils composed of collagen Types II and XI.

Type IX collagen is one of the more widely studied FACIT collagens. Its molecular structure contains helical features for integration into the fibril, as well as globular and armlike structures that allow interactions between fibrils. Immunochemical-labeling studies established that articular cartilage fibrils are biochemically heterogeneous, as different populations of fibrils share Type II collagen but have distinct compositions with respect to Type IX collagen, thereby defining their surface properties uniquely (Hagg *et al.*, 1998). Transgenic mice with mutations in a Type IX gene develop normally, but show degenerative changes in articular cartilage after birth. In addition, evidence for the importance of collagen IX in human articular cartilage comes from the recent finding that a mutation in one of the collagen IX genes causes multiple epiphyseal dysplasia (Olsen, 1997).

Collagen XVI is a minor FACIT component of fibrillar collagens (Kassner *et al.*, 2004). In cartilage, collagen XVI is a component of small heterotypic D-banded fibrils, mainly occurring in the territorial matrix of chondrocytes. Electron micrograph images of collagen XVI reveal rodlike molecules that harbor multiple sharp kinks. These kinks, possibly caused by noncollagenous regions, give rise to a highly flexible structure. The total length of individual trimeric recombinant collagen XVI molecules is about 240 nm, as calculated by atomic force and negative staining electron microscopy.

III. Fibrillar Molecular Packing

From the earlier work described here, it is clear that the collagen fibril is a complex structure. Many molecular species interplay to produce a fibril whose overall structure plays a specific role in the architecture of the tissue. In spite of such molecular heterogeneity, there are some central structural features of collagen packing that pertain more or less to all fibrillar structures. The purpose of this section is to discuss the generic and specific modulation in axial and lateral packing, and show how these lead to

reconciliation between observations relating to molecular crystallinity and overall fibrillar morphology.

The less contentious aspect of collagen molecular packing is the axial structure within a fibril. Fibril-forming collagens are typically 300 nm in length; the 67 nm density (D) step function repeat of fibrils, observed using many physical imaging techniques, is explained by the molecular stagger between molecules being 67 nm, or an integer multiple of this. Since the collagen molecular length is ~4.4 D, the molecular stagger leads in projection to regions of high and low electron density, these being the overlap and gap regions (Hodge and Petruska, 1963). The association between collagen molecules is driven by electrostatic and hydrophobic interactions, where a 234-amino-acid pseudoperiod observed within the collagen sequence of all fibril-forming types is the key to optimal electrostatic pairings between adjacent triple helices and maximizing the contact between hydrophobic regions (Hofmann *et al.*, 1980; Hulmes *et al.*, 1973; Itoh *et al.*, 1998; Ortolani *et al.*, 2000). The structural features of axial packing are shown in Fig. 1.

The axial association of heterotypic collagen molecules within a fibril is poorly understood. Indeed, the heterotypic axial register is relatively unstudied and information pertaining to interactions is mostly derived from crosslinking studies. This indicates a specificity of axial interaction between Type I collagen and Type III collagen. The interactions between Type I and Type V collagen are less well understood. Electron microscope evidence for the D-periodic interaction of Type IX collagen in a Type II-rich fibril was established by Vaughan *et al.* (1988). The basis of the interaction has been elaborated by biochemical studies that show Type II collagen forms extensive crosslinks with Type IX collagen (see, e.g., Eyre *et al.*, 2004). These interactions are further stabilized by the development of molecular crosslinks between the collagen molecules. These occur between sites in the short nonhelical N and C-terminal telopeptides of the collagen molecules and the main chain of the helix (for a review of collagen crosslinking see Bailey, 2001). Studies on oim/oim versus control tendons indicate that the total absence of alpha 2(I) chains results in a decreased order of axial packing and a loss of crystalline lateral packing. This suggests that the nonequivalence of three chains is an important determinant of lateral interactions between adjacent molecules and may be involved in the long-range axial order in Type I collagen-containing tissues (McBride *et al.*, 1997).

Detailed analysis of X-ray diffraction data and electron microscope data corresponding to the axial structure of fibrillar collagens has been made. These studies have indicated that the exact molecular periodicity contained 234.2 amino acids per D repeat (Meek *et al.*, 1979). The

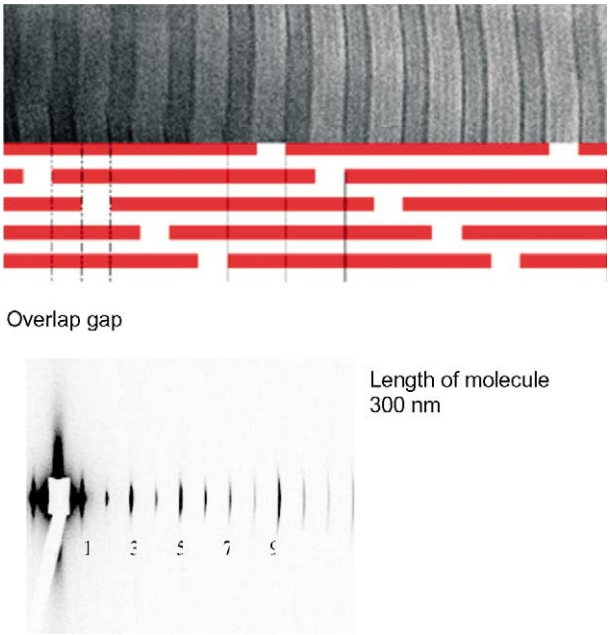


FIG. 1. The axial organization of collagen molecules in a collagen fibril. The pattern this arrangement produces is revealed by X-ray diffraction (bottom) and unstained cryoelectron microscopy (top). The individual 300 nm long collagen molecules are axially aligned in the fibril according to the Hodge and Petruska (1963) model (middle), where the collagen molecule's internal pseudo-periodicity facilitates staggered molecular interaction. This produces the gap-overlap step function of electron density that underlies the meridional series of reflections in the fiber diagram, and also produces the characteristic banding pattern of 67 nm seen in electron micrographs of collagen fibrils.

molecular organization of the telopeptide regions has been more difficult to ascertain. The majority of work has been carried out on the telopeptides from Type I collagen. Nuclear magnetic resonance (NMR) studies of the isolated telopeptides in solution yielded structural information indicating a contracted C-terminal peptide with a possible propensity to folding (Otter *et al.*, 1988). The modeling of isolated peptides using structure prediction models produced a variety of possible structures (see, e.g., Helseth and Veis, 1981; Helseth *et al.*, 1979; Jones and Miller, 1987). Such results may indicate that the structural stability of the telopeptides is conferred in part by the triple-helical environment in which they are surrounded in a fibril. Modeling studies of telopeptides constrained (or docked) within a hexagonally packed triple-helical environment have

been made by Jones and Miller (1991), in part by Vitagliano *et al.* (1995), and by Malone *et al.* (2004). These studies have indicated that the axial translation per residue within the telopeptides is less than that of the triple helix. This is in agreement with the interpretation of X-ray and neutron diffraction data made by Hulmes *et al.* (1977, 1980), where the N-terminal telopeptide sequence was shown to be contracted with a reduced axial rise per residue than seen in the triple helix. However, the C-terminal telopeptide contained a turn that causes the polypeptide chain to fold back on itself. X-ray diffraction studies by Bradshaw *et al.* (1989) also indicated that the position of heavy atoms labeling the tyrosine residues in the telopeptide required the C-terminal telopeptide to be folded. A profile of the electron density at a resolution of approximately 0.5 nm was obtained by Orgel *et al.* (2000). This indicated that the folding point of the Type I collagen $\alpha 1$ chain was between residues 13 and 14 of the C-terminal telopeptide. This feature also ensured that the telopeptide lysine residue involved in intermolecular crosslinking is in register with its acceptor sidechain on an adjacent helical segment. A schematic of this structure for the C-terminal region of the $\alpha 1$ chain of Type I collagen is shown in Fig. 2.

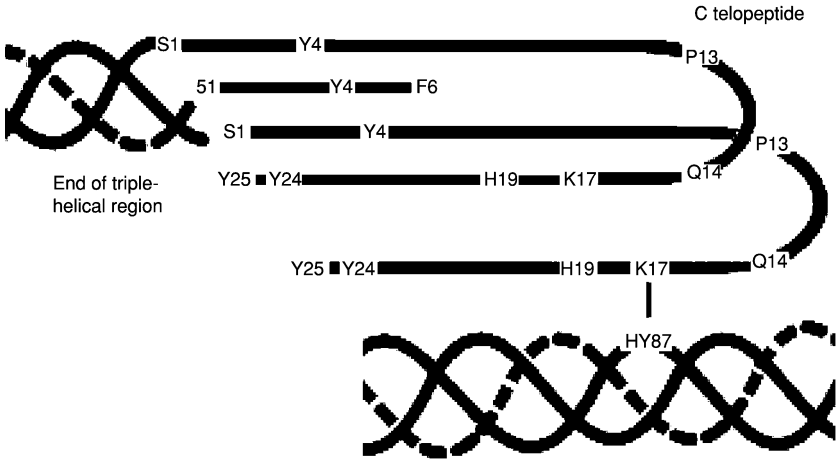


FIG. 2. Conformation of the C telopeptide. Determination of the projected electron density profile of the 67 nm repeat by employing conventional X-ray crystallography methods allowed the structure of the telopeptides to be shown. Here, a drawing of the C-telopeptide region shows a conformation that contains a reverse turn involving residues Pro13/Gln14. This would bring the Tyr24 residue into approximate axial alignment with the Tyr4 residues, and bring Lys17 into a favorable position to form the lysine–hydroxylysine cross-link with Hy87 and help stabilize the microfibrillar structure. Adapted with permission from Orgel *et al.* (2000).

The heterotrimeric structure of Types II and III collagen, and the lack of sequence similarity to Type I collagen, ensures that the conformation of telopeptides in a Type II or Type III collagen chain is different to that in Type I collagen. A study of the Type II collagen telopeptide structure utilized electron micrograph staining patterns that were analyzed and interpreted in terms of amino acid sequence (Ortolani *et al.*, 2000). A telopeptide model for both the N- and C-terminus was developed that contained molecular reversals at positions 10N–12N, 12C–14C, and 17C–19C for N- and C-telopeptides. This indicates that the telopeptides have an ''S-fold'' conformation, which can be interpreted as axial projections of a tridimensional conformation. In contrast, a study of the NMR solution structures of isolated telopeptides from Type II collagen (Liu *et al.*, 1993) indicated that the Type II C-terminal telopeptide is extended. The recent findings that Type II N-telopeptide interacts with Annexin V (Lucic *et al.*, 2003) points to a novel intercellular communication, which adds to the need for the local structure of this telopeptide region to be resolved.

The structures of the Type III and V collagen telopeptides have been less studied. However, the NMR study of Type III telopeptides has been reported, and the 22-amino-acid C-terminal telopeptide is extended with a tight turn involving residues 8–11 (Liu *et al.*, 1993). Crosslink analysis reveals connectivity between the C-terminal telopeptide of Type III collagen and the N-terminal helical region of another Type III molecule (Henkel, 1996).

The telopeptides are characterized by a lack of the classical triplet Gly-Pro-Hyp. This characteristic signature is the most frequently occurring sequence in fibril-forming collagens, but it only accounts for about 10% of the amino acids triplets found. The significance of sequences with low Hyp/Pro occupancy have led to suggestions that real collagen sequences may contain local distortions from the accepted helical structure formed by ''model'' Gly-Pro-Pro peptides (Paterlini *et al.*, 1995). Simple modeling of the electron density profile of collagen defined by the presence of amino acids at discrete locations produces an adequate fit with electron micrograph data and X-ray diffraction data. However, the expected position of amino acids requires fine tuning in order for an optimal fit with the data to be obtained (Brown *et al.*, 1997; Orgel *et al.*, 2000). This offers the possibility of regions of rarefaction and relative compression in the axial density profile. This may in part explain why crystallization of real collagen sequences has proven difficult. The presence of local distortions in the structure may be induced by the fibril formation; these areas could correspond to regions of differing molecular extensibility, thereby bestowing unique mechanical properties to the fibril (Silver *et al.*, 2001).

IV. Lateral Packing in Fibrillar Collagens

Agreement on the basis for lateral packing of collagen molecules within a fibril is less well resolved than for axial packing. Some fibrillar structures, such as those seen in rat tail tendon, have attracted particular attention as they exhibit long range crystallinity (North *et al.*, 1954; Wess *et al.*, 1998), as do turkey tendon (Jesoir *et al.*, 1981) and lamprey notochord, a Type II fibrillar structure (Eikenberry *et al.*, 1984). Such structures have provided a basis for study since long-range order can provide a rich source of data, giving indications to a commonality in molecular packing that may pervade all collagen fibrillar structures.

Detailed investigations of packing modalities within fibrils from rat tail tendon favor the presence of microfibrillar structural units that are arranged as compressed microfibrils on a triclinic unit cell lattice (Fraser *et al.*, 1983; Wess *et al.*, 1995). Here, the pentameric topology suggested by Smith (1968) is adhered to, while a more realistic packing density is achieved by distortion of a regular pentagon to a compressed structure that allows quasi-hexagonal lateral packing (Trus and Piez, 1980). These units have been revealed in more detail recently by the determination of the structure of a single collagen unit cell from the electron density map derived from X-ray diffraction studies. This was achieved using adapted protocols from the multiple isomorphous replacement technique of phase determination in macromolecular crystallography (Orgel *et al.*, 2001). This study confirmed previous interpretations of the fiber diagram, where the main structural features were kinked collagen molecules tilted relative to the fibril axis.

A feature gleaned from earlier interpretation of the quasi-hexagonal molecular packing scheme for collagen was that the molecules within the overlap region have a common direction of tilt of magnitude ~5 degrees, with a direction that is almost parallel with the *b* axis of the triclinic unit cell (Miller and Tochetti, 1981). In order to ensure that the contents of all unit cells are identical in a crystallite, there must be a rearrangement of molecular connectivity in the gap region. This implies that one collagen molecule contains several kinks, where the overall magnitude of its tilt is similar in both the gap and overlap regions. However, the azimuthal orientation of the tilt varies to ensure that each D segment passes through the correct topological position of the unit cell.

A consensus of opinion from these studies points to a one-dimensional staggered microfibrillar structure with intermicrofibrillar crosslinks. Such intermicrofibrillar links are important in the hierarchical connectivity at the supramicrofibrillar level and also provide the basis as to why individual microfibrils have proven so difficult to isolate. The molecular topology of

compressed microfibrillar structures and three-dimensional electron density maps are shown in Fig. 3. Evidence for a microfibrillar structure has also been obtained using a number of other physical characterization techniques. For example, electron tomographic reconstructions in dry cornea fibrils revealed ~4 nm microfibrillar-type structures (Holmes *et al.*, 2001). Atomic force microscopy (AFM) was used to reveal microfibrils of Type I collagen arranged parallel, or inclined approximately 5 degrees, to the fibril axis (Baselt *et al.*, 1993).

In spite of evidence for specificity of molecular packing indicated by crystallinity, the X-ray fiber diagram of fibrillar collagen also contains a significant amount of diffuse scatter, indicating that fibrils contain a large amount of static or even dynamic disorder. Fibrils that contain crystalline regions are also thought to contain significant levels of collagen molecules exhibiting liquid-like disorder, where the lower density gap region of the fibril structure is believed to be more disordered than the overlap region (Fraser *et al.*, 1987; Wess *et al.*, 1998).

Indeed, the molecular packing of collagen has been likened to that of a liquid or liquid crystal, where only local molecular interactions are of significance (Fratzl *et al.*, 1993; Hukins and Woodhead–Galloway, 1977; Knight and Vollrath, 2002). In some fibril structures, this type of scatter is the only feature observed by X-ray diffraction, and the overall molecular packing can be described as liquid-like. These observations point to a variety of levels of lateral molecular organization in collagen fibrils, ranging from liquid-like to crystalline.

The relationship between crystallinity, disorder, and supramolecular topology of crystalline packing within a fibril requires explanation. Organized lateral packing within fibrils results in a long-range crystalline structure, and this has been observed in lateral sections of fibrils studied using electron microscopy (Hulmes *et al.*, 1985). A further investigation by Hulmes *et al.* (1995) attempted to assess the compatibility of the multitude of proposed fibrillar packing models that had emerged with electron micrograph and X-ray diffraction data. Models of molecular packing within fibrils, ranging from ordered crystalline to disordered liquid-like structures, were constructed and their corresponding Fourier transforms compared to the diffraction pattern of tendon. All of the models studied were found to be unable to accommodate both crystallinity and disorder to the appropriate extent, or else contained unrepresentative molecular correlations. A model developed within this particular study based on concentric rings of microfibrillar structures was found to have the best agreement with the diffraction data. The structure contained sectors of crystalline order interfaced by disordered grain boundaries; this allowed crystallites to be accommodated systematically within a fibril with a circular

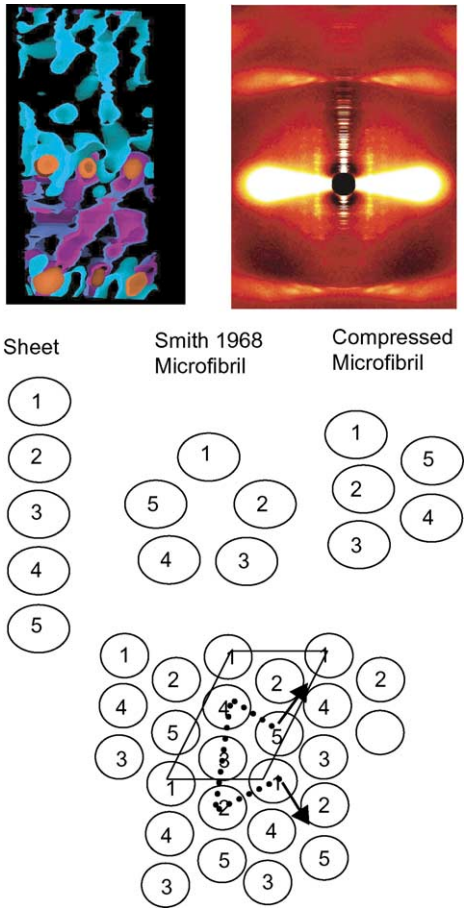


FIG. 3. X-ray diffraction pattern of collagen fibrils in rat tail tendon. The X-ray diffraction pattern of rat tail tendon (top right) contains a series of intense meridional reflections that relate to the axial molecular organization. Three-dimensional microcrystalline domains of collagen molecules produce discrete Bragg peaks in the equatorial (m = 0; n = 0) helix layer plane and parallel to the meridian of the diffraction pattern. The resolution of the diffraction data extends to approximately 1.0 nm along the equator and 0.54 nm parallel to the meridian. The Bragg diffraction peaks from such images were used to determine the electron density within a unit cell adapted with permission from Orgel *et al.* (2001) (top left). This clearly shows the molecular path of individual helices passing through the unit cell. Here, the long *c* axis has been compressed by a factor of eight in order for the molecular path to be seen. Many different models have been proposed to account for the 3D packing of collagen molecules in the type I fibril. One of the earliest was based on a five-stranded pentagonal microfibril, where the molecular translations were cyclic. This was later replaced by compressed microfibrillar structures that pack together to produce a crystalline array. The most likely topology of interaction determined by Wess *et al.* (1998) is also shown as a dotted line set in a crystalline array of compressed microfibrils. The arrows indicate the position of the telopeptide connections between microfibrils.

cross-section. The organization of the crystallites, with the *a* axis of the lateral unit cell aligned in a radial direction and the *b* axis tangential to the fibril surface, may explain the evidence for curvature of the crystallites described by Fraser *et al.* (1983). The model also provided a structural basis for growth with the incremental deposition of molecules or microfibrils at the fibrillar surface, as suggested by the 8 nm quantal variation in fibril diameter observed by Parry and Craig (1979). The cross-sectional structure of the proposed fibrillar structure is shown in Fig. 4.

The molecular tilt observed in tail tendon structures is about 5 degrees relative to the fibril axis. It has been proposed by Ottani *et al.* (2001) that such a tilt is typical of a structural class of fibril found in tissues that resist uniaxial stress, such as tendon, ligament, and bone. A distinct class of fibril are those that exhibit a molecular tilt of greater magnitude, ~18 degrees. This can be seen in the fibrils from tissues such as skin (Brodsky *et al.*, 1980), chordae tendinae (Folkhard *et al.*, 1987a), cornea (Marchini *et al.*, 1986; Yamamoto *et al.*, 2000a), blood vessels, nerve sheaths (Ottani *et al.*, 2001), and submucosa (Cameron *et al.*, 2002). In these tissues, fibril diameters are relatively uniform, typically small (less than 100 nm), and the overall mechanical requisite of the tissue is to resist multidirectional stresses. This leads to a characteristic shortening of the axial repeat from 67 to 65 nm, where 65 equals 67 cos 18°. Evidence from X-ray diffraction and electron

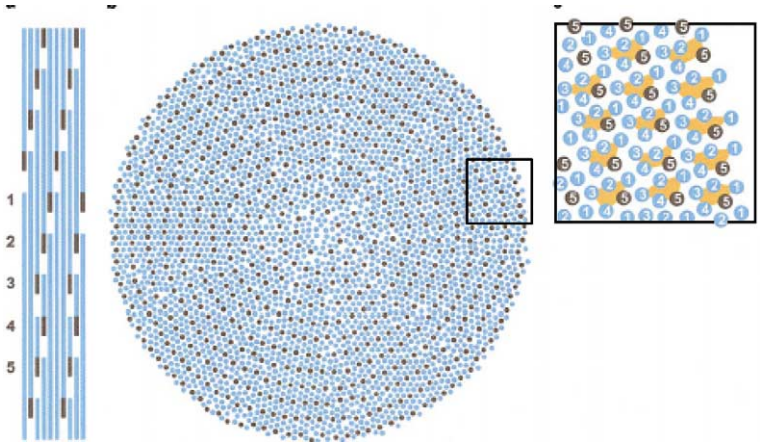


FIG. 4. A model for the possible relationship between crystalline and disordered regions within a collagen fibril. The cross-sectional model of a 50-nm diameter fibril shows regions of crystallinity interfaced by grain boundaries. The individual crystalline unit cells are shown and the gap region is represented by a darker color. The axial projection of a single microfibrillar unit is also shown. Based on the structures developed by Hulmes *et al.* (1995) and adapted with permission from Hulmes *et al.* (2002).

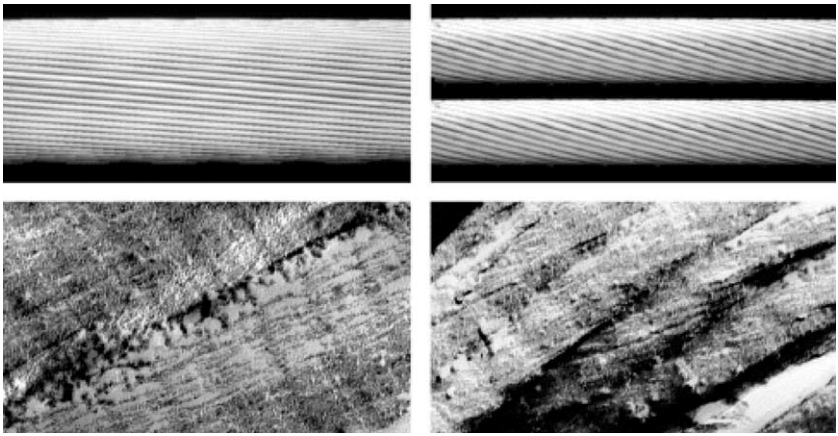


FIG. 5. Freeze etched micrographs and molecular models of collagen fibrils exhibiting straight and helicoidal conformations. Molecular models of the subfibrillar structures in parallel arrangement (top left) and fibrils with helical arrangement (top right). Below, the same arrangements shown by freeze-fracture micrographs of corresponding hydrated, unfixed rat tail tendon (left), and bovine cornea (right). Tail tendon consists of slender subfibrillar structures winding at an extremely shallow angle, while the subfibrillar structures that comprise the collagen fibrils of cornea wind at an angle of ~17 degrees with respect to the fibril axis. Adapted with permission from Ottani *et al.* (2001).

microscopy shows that the molecular tilt of the molecules is believed to be constant throughout the fibrils (Holmes *et al.*, 2001). This implies that such helicoidal arrangements must lead to the disruption of intermolecular or intermicrofibrillar interactions. An electron micrograph and model of helicoid and ''straight'' fibrils are shown in Fig. 5. The basis for helicoid fibril formation is not clearly defined, although differences in the helical lengths of heterotypic fibril-forming collagens could play a part in defining the helical path of molecules within a fibril (Cameron *et al.*, 2002). These tissues also typically demonstrate a lack of crystallinity and a more liquid-like packing. It is attractive to speculate that the minor molecular species interrupt the crystalline packing of the major collagen species, but this remains to be proven.

V. Novel Fibrillar Structures

Up to this point, all collagen fibrillar structures discussed are of the 65 to 67 nm axial repeat type, the most prevalent collagen form in animals. However, additional forms have been observed *in vivo*, or in pathological

or novel tissues. These also deserve some attention, since they show the further diversity of collagen fibrillar structure. Fibrous long spacing collagen (FLS) fibrils (Highberger *et al.*, 1950) are collagen fibrils that display a banding with periodicity greater than the 67 nm periodicity of native collagen. Typical periodicities range from 150 to 250 nm. FLS collagen has been described both *in vitro* and *in vivo*. An electron micrograph and putative model of the structure is shown in Fig. 6. In particular, FLS collagen has been found both in pathological and in normal tissues (Kamiyama, 1982; Morris *et al.*, 1978; Nakanishi *et al.*, 1981; Slavin *et al.*, 1985). FLS may be the result of partial degradation of collagen reticular fibrils by endogenous collagenase, or possibly a derivation from the association of immature collagen microfibrils and acid mucopolysaccharides present in tumoral tissues. FLS fibrils can be formed *in vitro* by addition of oil-acid glycoprotein to an acidified solution of monomeric collagen, followed by dialysis of the resulting mixture. The assembly and formation of FLS and the characteristic banding pattern of the fibrils are discussed in Paige *et al.* (2001).

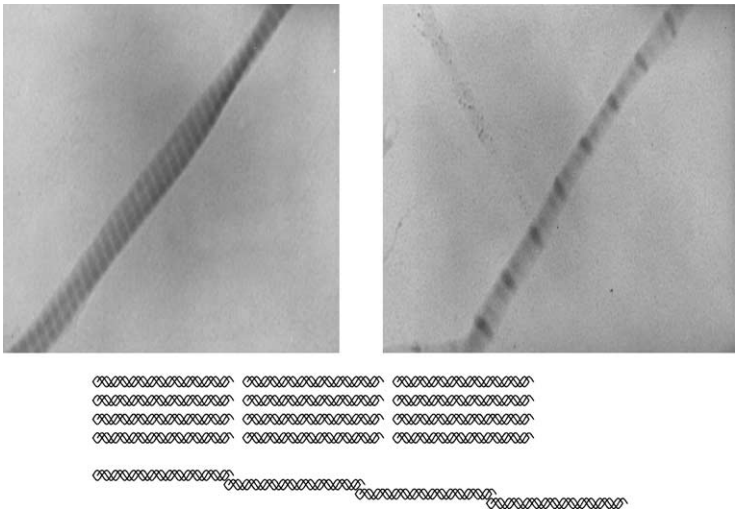


FIG. 6. An electron micrograph of normal banded (left) and filament long spacing (FLS) collagen (right). The proposed relationship of the collagen molecules with the extended 250 nm spacing corresponds to approximately 4D where D = 67 nm spacing. The exact molecular basis for the periodicity remains unresolved. Possible end-to-end molecular association or 4D stagger are shown below. These probably would require reinforcement by ''scaffolding'' molecules, such as α1 glycoprotein. Figure provided by K. Smith, University of Strathclyde.

Other periodic fibrillar structures with D periodicities less than 67 nm have been reported on a number of occasions. Fibrils exhibiting banding patterns with 23 ± 2 nm periodicities have been described by a number of researchers (Porter and Pappas, 1958; Venturoni *et al.*, 2003), and a 9.0 nm staggered microfibrillar structure was reprecipitated from solution (Doyle *et al.*, 1974). Industrial processing, such as dehydrothermal heating, can also produce variants in fibrillar structure (Gorham *et al.*, 1991).

VI. Subfibrillar Structures

There is strong evidence that certain fibril types contain subdomains of structure that lie between molecular/microfibrillar and fibrillar levels. Such nanoassemblies are of great interest, since they may reveal important information that relates to the overall mechanical properties of the fibrils and also the nucleation and growth processes of fibril formation. Recent AFM studies by Wen and Goh (2004) reveal fibrillar substructures, and work by Gutsmann *et al.* (2003) led to the concept that the collagen fibril is an inhomogeneous tubelike structure composed of a relatively hard shell and a softer, less dense core. It should be noted that such density differences may be possible from the packing model proposed by Hulmes *et al.* (1995). In contrast, evidence from Franc (1993) showed central condensed material of 4 nm diameter in all collagen fibrils. This persisted even when fibrils had been swollen and partly disorganized by ethyl glycol dehydration. Further cytochemical characterization indicated the glycoprotein nature of this central compact material. The case for a generic fibrillar substructure therefore remains unresolved, and it is possible that bespoke accretion properties may be directed by fundamentally different subfibrillar architectures.

VII. Fibril Surface

The surface of the fibril provides the interface between the internal structural and mechanical properties of a fibril, and the rest of the extracellular matrix. The surface, therefore, is the most complex area of the fibril in terms of molecular heterogeneity and structure. Caution has to be taken in the comparison of structural and biochemical evidence from complementary techniques since the extraction, dehydration, and sample preparation can cause variation in observations of fibril surface properties (Raspanti *et al.*, 1996).

The distribution of minor collagen types as surface molecules is more obvious for some types than others. Fibril-forming collagens—especially

Types III, V, and XI—may decorate fibril surfaces with their partially processed N-terminal regions. A possible role of these moieties in the regulation of fibril diameter is discussed later.

The FACIT collagen types have structural features that would clearly disrupt the internal structure of a collagen fibril, and the electron micrographs demonstrating Type IX collagen protruding from the surface of fibrils establish this class of molecules as modulators of surface properties. The N-terminal NC4 domain of Type IX collagen is a globular structure projecting away from the surface of the cartilage collagen fibril. Several interactions have been suggested for this domain, reflecting its location and its characteristic high isoelectric point (Pihlajamaa *et al.*, 2004). Binding assays showed that the NC4 domain of Type IX collagen specifically binds heparin at a site located in the extreme N terminus containing a heparin-binding consensus sequence, whereas electron microscopy suggested the presence of at least three additional heparin-binding sites on full-length Type IX collagen. The NC4 domain was also shown to bind cartilage oligomeric matrix protein. Type IX collagen appears to shield Type II collagen from exposure on the fibril surface. This is probably due to the presence of chondroitin sulphate sidechains, which will also play an important role in maintaining fibril spacing (Bishop, 2000). A model of Type IX distribution on the fibril surface proposed by Eyre *et al.* (2004) can accommodate potential interfibrillar as well as intrafibrillar links between the Type IX collagen molecules themselves, so providing a mechanism whereby Type IX collagen can stabilize a collagen fibril network. Such links at the interfibrillar level are essential for the maintenance of structural integrity.

In addition to a number of collagen gene products crowding the fibril surface, contributions are made by several noncollagenous proteins. Principally, these are the proteoglycans such as decorin, lumican, and fibromodulin. Their role is thought to involve intrafibrillar connectivity and matrix-cell communication. They are also essential modulators of the interfacial shear between fibrils. Most proteoglycans show a nonrandom axial distribution, with a clear preference for the gap region within the D-period (see, e.g., Danielson *et al.*, 1999; Hedlund *et al.*, 1994; Scott and Thomlinson, 1998).

VIII. Factors that May Regulate Fibril Growth and Size

Since fibril assembly can be regarded in part as a spontaneous self-assembly process, the limitation of fibril size could be ascribed to a physical equilibrium between soluble procollagen molecules and the growing insoluble fibril. Fibril-forming collagens are synthesized as precursor

procollagens, where N- and C-terminal globular propeptide extensions maintain solubility. The C-propeptide directs chain association during intracellular assembly of the pro-collagen molecule from its three constituent alpha chains. During secretion and deposition as the extracellular matrix, the globular propeptides are cleaved by specific procollagen proteinases, triggering fibril formation (Prockop and Hulmes, 1994). Certain accretion models have postulated diffusion-limited models for simple collagen systems (Parkinson *et al.*, 1995), and growth by this principle would result in discrete fibrils of relatively fixed diameter. However, tissues such as tendon, which are predominantly Type I collagen, seem to contain a wide variety of fibril dimensions. This may indicate a lack of regulation in fibril diameter by a simple diffusive model, and other factors may be important here. In the extracellular matrix, the procollagen C-propeptides ensure procollagen solubility, while the persistence of the N-propeptides controls fibril shape (Hulmes, 2002). It has also been suggested that the N-terminal propeptides that persist to different extents in different collagen chain types play a role in determining fibril size distribution. Persistence of the N-propeptide from the slower processing of Type III collagen or the partial removal of the bulky propeptide units of Types V or XI collagen may result in a decoration of the fibril surface with globular protein domains. This may prevent further accretion, and various possibilities are discussed by Chapman (1989), Birk (2001), and Linsenmayer *et al.* (1993). Such surface interactions may be an important factor in why many fibrils are heterotypic in nature.

Type V collagen integration in corneal tissue from both avian and mammalian sources points to heterotypic collagen interactions providing some basis for fibril diameter regulation (White *et al.*, 1997). Indeed, the triple-helical structure of Type V collagen alone has been shown to be an important factor in regulating the fibril diameter of Type V/Type I heterotypic fibrils *in vitro* (Adachi and Hayashi, 1986).

Proteoglycans also have an important role in the regulation of fibril structures. Molecules such as lumican (Chakravarti *et al.*, 1998), decorin (Danielson *et al.*, 1999), and fibromodulin (Svensson *et al.*, 1999) coat the surface of fibrils in a specific manner and may restrict further growth. The switch mechanisms that may be responsible for the coating of fibrils after nucleation or a certain amount of growth are still subject to speculation (MacBeath *et al.*, 1993). Knockout gene experiments in mice have indicated that the loss of proteoglycans (such as decorin) causes alteration of fibrillar morphology, and fibril sizes are observed to be more polydisperse. However, upregulation and compensatory behavior of other associated proteoglycans make the results more complex to interpret than simply the mere removal of a specific gene product (Derwin *et al.*, 2001).

Other knockout experiments designed to investigate fibril size in SPARC-(secreted protein acidic and rich in cysteine) null skin detected that the fibrils were smaller and more uniform in diameter in comparison to those of wild-type skin (Bradshaw *et al.*, 2003). At five months of age, the average fibril diameter in SPARC-null versus wild-type mouse skin was 60.2 nm versus 87.9 nm, respectively. Analysis of the extractable collagen species indicated a relative increase in Type VI collagen, accompanied by a decrease in Type I collagen. In SPARC-null mice, possible interactions between Types I and VI collagen have been described, although they are poorly understood.

Work by Haston *et al.* (2002) also indicates that acidic glycoprotein (AGP) influences Type II collagen fibrillogenesis, where *in vitro* studies show that low concentrations of AGP produced decreases in fibrillogenesis rate and fibril diameter. High concentrations produced fibrils at a rate and diameter dependent on fucosylation of AGP. Highly fucosylated AGP produced narrow fibrils, and poorly fucosylated AGP produced thicker fibrils.

The interplay of heterotypic collagens, fibril-associated collagens, and cofibrillar macromolecules that alter the accretion properties of available procollagen molecules may be required to form a typical collagen fibril with its characteristic cylindrical central shaft section and differential tip shapes. Although collagen fibril formation can be mimicked *in vitro* as an acellular system, the role of cellular processes cannot be underestimated. The challenge remains in research to understand the interplay that exists between extracellular macromolecules and cellular processes that produce the final fibrillar structure.

IX. Distribution of Fibril Sizes

Fibril-forming collagens are typically 300 nm long and approximately 1 nm wide. Such regular structures, however, form fibrillar structures of length and diameter that tend to vary depending on anatomical location. The basis for this wide distribution of fibrillar architectures is not well understood, but may lie in the composite nature of many fibrils. The effective length of fibrils has been difficult to study, since *in vitro* fibril preparation is probably not reflective of the overall length of a fibril *in vivo*. Furthermore, the lengths of the fibrils probably change with maturation. Disruption and dispersion of fibril structures from tissues inevitably leads to breakage of fibrils, and resultant intact length may not be representative of the overall fibrillar length. Some attempts have been made to study fibril lengths in electron microscope transverse section. This is a painstaking and time-consuming pursuit; therefore, the number of fibrils that can

be studied systematically in a tissue is low. The study by Craig *et al.* (1989) estimated the length scale of rat tail tendon fibrils to be from 0.3 mm to greater than 10 mm, depending heavily on developmental status. Studies of fibril length in development by Birk *et al.* (1997) failed to find both ends of any fibril. This indicated a possible fibril length longer than sequential microscope sections would allow, or that the fusion of fibrils produced a networked continuum at maturity. A comparison of fibril length and diameter measurements using different approaches is described in Redaelli *et al.* (2003).

Collagen fibril diameters vary over at least two decades of nanometer length scales, but the reasons for such a large variation amongst tissues remains poorly understood. It has been suggested that the mechanical properties of tendon are related to the fibril diameter distribution; the large fibrils have a primary role in withstanding high tensile forces and the smaller fibrils have a special ability to resist creep (Parry and Craig, 1984). Fibril diameters increase with maturation of tissue, but break down at senescence (Parry *et al.*, 1978). However, experiments monitoring changes in fibril diameter with *in vivo* loading of tissue have resulted in little consensus (Michna, 1984; Patterson-Kane *et al.*, 1997). Tissues such as mature Achilles tendon contain a bimodal distribution of diameters with thicker fibrils of diameters about 150–250 nm and a population of thinner fibrils with diameters of about 50–80 nm (Svensson *et al.*, 1999). In structures such as cartilage and vitreous humor, there is also a clear bimodal distribution of fibril diameters; however, many tissues such as skin and submucosa contain a more uniform distribution of diameters around 60 nm (Sanders and Goldstein, 2001). The rationale behind a discrete fibril diameter in load-bearing tissues is not immediately apparent, since *a priori* a distribution of fibril sizes would be thought to dissipate the load more evenly throughout the tissue. It has been argued that these tissues are usually associated with tear resistance and higher compliance than either tendon or ligament. The fibril diameter distribution in a tissue is possibly, therefore, a balance between overall tensile strength and creep resistance, the former being related to fibril size and the latter to the interfacial shear and thus the surface area:volume ratio (for a review, see Ottani *et al.*, 2001). A special case where the optical properties of the tissue are paramount is that of cornea. Probably the most highly regulated fibrillar diameter in animals is that found in adult human cornea (~33 nm; Meek and Fulwood, 2001). The relatively small and highly constrained diameters, and the distribution on a semicrystalline lattice, are both essential features for the transparency of the tissue.

Care has to be taken in comparing fibril diameter information from a variety of sources, since sample preparation and the ability of the

technique to estimate the distribution of fibril sizes may result in different fibril size distributions being obtained. In particular, standard embedding techniques for electron microscopy led to an underestimation of fibrillar diameter by some 33% due to sample shrinkage on preparation. Techniques such as small-angle X-ray scattering, which has been used to estimate fibril diameter without embedding or sectioning, tends to skew the distribution to larger fibril sizes unless the scattering power from large fibrils up to ~500 nm in diameter is taken into account.

X. Suprafibrillar Architectures

Although there is a strong molecular and nanoscopic driving force behind our understanding of the functionality of collagen fibrillar structures, the mesoscopic and macroscopic properties of the tissue are where the structural integrity manifests itself. One of the differences in mechanical resistive properties between tissues such as skin and tendon is the feltwork nature of fibril distribution in skin. This allows resistance to strain to occur within a two-dimensional plane. In contrast, tendon is required (normally) to resist strain only along its axis.

Cornea is a tissue where collagen fibrils of uniform diameter are regularly arranged within a matrix of proteoglycans. The fiber direction in cornea, however, is not uniform and consists of a succession of superposed layers of fibrils with different orientations. The collagen fibrils in each layer are parallel to one another, but are kept at a significant distance from one another. This is essential for the transparency of the tissue. Fibril distribution in the cellular cementum present a lamellar packing pattern that conforms to the twisted plywood principle of bone lamellation, with a periodic rotation of matrix fibrils resulting in an alternating lamellar pattern (Yamamoto *et al.*, 2000b). In normal articular cartilage, the collagen fibrils in the superficial zone are compactly arranged into layers of decussating flat ribbons, mostly parallel to the artificial split lines used for specimen orientation (Hwang *et al.*, 1992).

Collagen fibrils associate in many tissues to form discrete fiber structures. Within one collagen fiber, the fibrils are oriented not only longitudinally but also transversely and horizontally. The longitudinal fibers do not run only parallel but also cross each other, forming spirals, and some of the individual fibrils and fibril groups form spiral-type plaits (Kannus, 2000). Such local suprafibrillar architectures present organizational features of lamellae, twisting or sinusoidal ''crimps'' on a mesoscopic scale. Examples of these structures can be seen in Fig. 7. It has been postulated that a possible commonality between many different collagen-based tissues

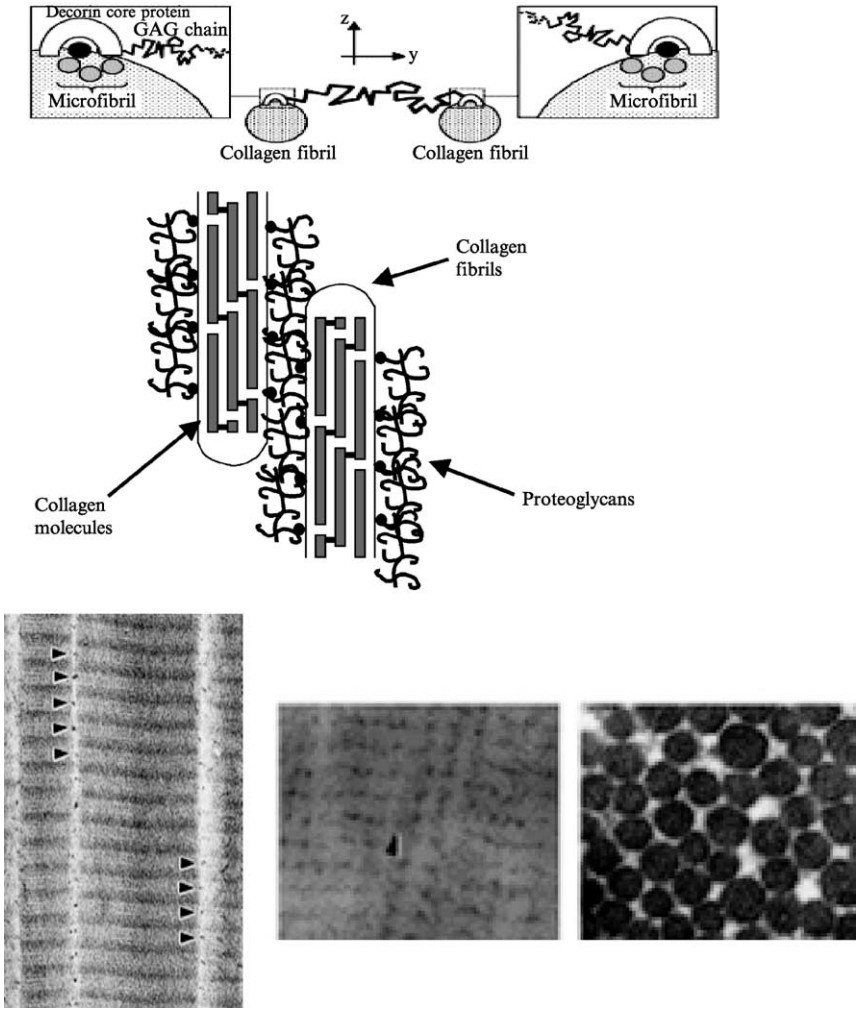


FIG. 7. The collagen fibril surface, which is an area important for interfacial interactions. Electron micrograph image of the collagen fibril in longitudinal and transverse section, adapted with permission from Kuwaba *et al.* (2002). In both cases, the proteoglycan staining follows the 67 nm periodicity, where metal ion staining of the proteoglycan reveals the presence of proteoglycans such as decorin. The drawing shows the possible way in which proteoglycan GAG chains provide the interfacial shear between collagen fibrils (adapted with permission from Fratzl, 2003). The possible connectivity between two microfibrils in opposing fibrils is shown in the figure adapted with permission from Vesentini *et al.* (2004).

is that the suprafibrillar organization resembles the cholesteric or pre-cholesteric organization of liquid crystals (Giraud-Guille, 1996; Hulmes, 2002). Such liquid crystalline behavior was shown to occur in very high concentrations of collagen at low pH; this behavior has also been observed in procollagen molecules at high concentrations in a physiological buffer (Martin *et al.*, 2000). Therefore, it may be possible that liquid crystalline ordering of collagen may precede condensation into true fibrillar forms.

Relatively few observations have been made that relate to the specificity of intrafibrillar interactions. In the main observation, fibrils are not required to have a regular axial spatial relationship that would lead to long-range coherence in the tissue. Mechanical models of interfibrillar relationships usually assume nonspecific interactions occurring through the proteoglycan-rich gel and hydrated milieu between fibrils. In the vitreous humor, thin heterotypic collagen fibrils have a coating of non-covalently bound macromolecules which, along with the surface features of the collagen fibrils themselves, probably play a fundamental role in maintaining gel stability. They are likely to both maintain the short-range spacing of vitreous collagen fibrils and to link the fibrils together to form a network. A collagen fibril-associated macromolecule that may contribute to the maintenance of short-range spacing is opticin, a leucine-rich repeat protein. Collagen fibrils in extracellular matrices of connective tissues (tendon, cornea, etc.) are also bridged and linked by the anionic glycosaminoglycans (AGAGs) of the small proteoglycans (decorin, etc.). Such bridges maintain the collagen fibril supramolecular architecture and were proposed as *shape modules* by Scott and Thomlinson (1998). Figure 8 shows an electron micrograph and schematic of the fibril surface. The close association of proteoglycans or glycosaminoglycans (PGs/GAGs) with maturing Type VI collagens may provide a further indication to the link that associates D-periodic collagen fibrils via PGs/GAGs. Ruthenium red-stainability on the surface of D-periodic collagen fibrils was also examined, the results showing that these sites were also D-periodically associated (Watanabe *et al.*, 1997).

Furthermore, specific interactions of collagen fibrils with the GAG-rich regions of several aggrecan monomers aligned within a proteoglycan aggregate have been identified. The fibril could therefore serve as a backbone in at least some of the aggrecan complexes (Hedlund *et al.*, 1999). A highly specific pattern of crosslinking sites suggests that collagen Type IX has evolved to function as an interfibrillar network-bonding agent. Interfibrillar coherence has been observed in some mineralized tissues, such as turkey leg tendon (Landis *et al.*, 1996), where the axial banding pattern of several mineralized fibrils is in register. This may result

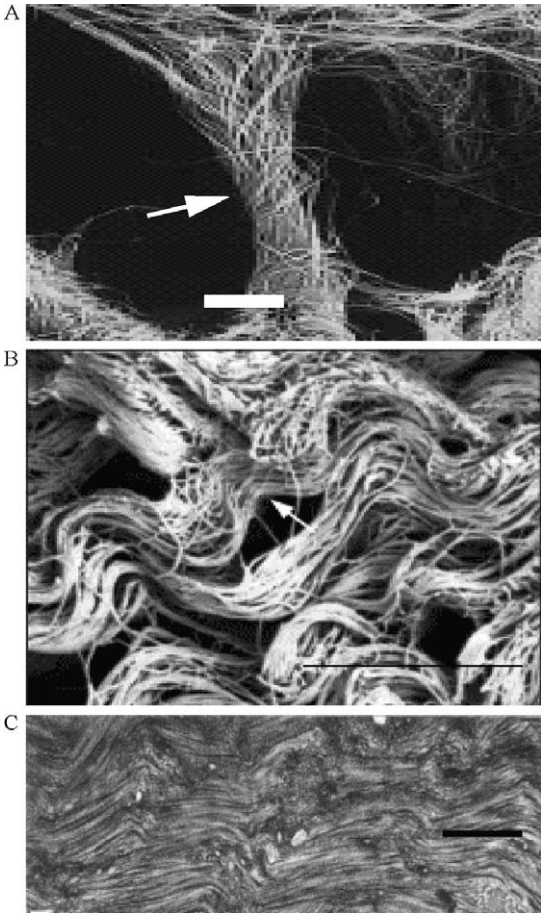


FIG. 8. Suprafibrillar architectures reveal a number of features that resemble cholesteric and precholesteric states. (A) Fibrils at a lesion between a bone-tendon interface imaged by SEM show the intertwining helical nature of fibrillar interactions. Adapted with permission from Oguma *et al.* (2001). (B) Scanning electron microscope views of the collagen fibrillar network of the mouse pubic symphysis. Adapted with permission from Pinheiro *et al.* (2004). (C) Cross-striated collagen fibrils in a stabilized collagen gel after neutralization. The fibrils follow undulating patterns and mimic geometries described as crimp morphologies in tendons (bar = 0.5 μm). Adapted with permission from Giraud-Gille *et al.* (2003).

from the growth of mineral between fibrils, but may be an association required prior to interfibrillar calcification. The significance of this, and the specificity of interfibrillar interactions, remains a rich vein for future research.

The distribution of fibrils as fibers or fibril bundles occurs as a common motif in many tissues, with the fiber direction being aligned to provide the bespoke mechanical characteristics of the tissue. Suprabundle structures of collagen often provide the interface of collagen with other biopolymers, such as elastin and fibrillin. These elastic components complement the stiff collagen fibers and produce the overall functional properties of the tissue. Periodontal ligament contains principal fibers developed from aggregates of fibrils that are allowed to form during root development in the space between ligamental cells (Yamamoto and Wakita, 1992). In the case of tendon, the suprafibrillar hierarchical structures are well characterized. Here a group of collagen fibrils forms a collagen fiber—the basic unit of a tendon. Fibers are bound together by a fine sheath of connective tissue called the endotenon. A bunch of collagen fibers forms a primary fiber bundle, and a grouping of primary fiber bundles results in a secondary fiber bundle. Secondary fiber bundles accumulate to form tertiary bundles that comprise the tendon (Kannus, 2000).

XI. Mechnical Properties of Collagen-Rich Tissues

The outstanding mechanical properties of tissues that rely on fibrillar collagen are due to the optimization of their structure on many hierarchical levels. The interplay between structures already described from the molecular to fibrillar, and thence to interfibrillar, are critical to the overall mechanical properties (Fratzl, 2003). Strain in the tissue results from two principal mechanisms—the molecular elongation and molecular shear within a fibril, and the shear deformation of the proteoglycan-rich matrix between fibrils. The correct supramolecular organization and crosslinking within each fibril is therefore essential. The response of the tissue to the stress imposed depends very much upon the strain rate, where high strain rates cause elongation of the collagen molecules and initiate shearing effects within a fibril. Slow shear rates result in shear of the matrix between the fibrils, leading to creep (Sasaki *et al.*, 1999).

Synchrotron radiation studies of fibril behavior in tissues have been critical to investigations of structural alterations, since they allow transient structural features to be monitored in realistic time frames. Stress-strain experiments conducted synchronously with X-ray diffraction reveal a fiber indicating that only 40% of the increase in the D period before breakage—from 67 nm to just under 69 nm—is from the extension of the collagen helix. The rest of the fibril extension must result from some form of molecular rearrangement. This appears with the change in the relative intensity of the second- and third-order Bragg peaks of the meridional series, indicating a severe distortion in the usual gap/overlap step function

that is characteristic of fibrillar collagens (Sasaki and Oadjima, 1996). *In situ* synchrotron X-ray scattering experiments suggest that several different processes could dominate, depending on the amount of strain. While at small strains there is a straightening of kinks in the collagen structure (first at the fibrillar and then at the molecular level), higher strains are believed to lead to molecular gliding within the fibrils and ultimately to a disruption of the fibril structure. Preliminary attempts to quantify these distortions in terms of alteration of the collagen axial structure met with limited success (Folkhard *et al.*, 1987b). Here, in similar studies, three elementary models for molecular elongation and rearrangement of collagen within a fibril were proposed. In the first model, molecular elongation arises through an alteration of the helical pitch. In the second model, there is an increase in the length of the gap region. The third model shows that a relative slippage of laterally adjoining molecules occurs.

The shortcomings of such investigations can be interpreted as: (a) the starting fibrillar model is of insufficient accuracy to mimic the changes imposed by the mechanisms of distortion; or (b) the emphasis, or even the entire basis, for the mechanisms of distortion is incorrect.

It has also been noted that the strain within collagen fibrils is always considerably smaller than in the whole tendon. This phenomenon is still very poorly understood, but points toward the existence of additional gliding processes occurring at the interfibrillar level (Fratzl *et al.*, 1998). Once again, this leads to the fibril surface and interfibrillar interactions having a more prominent role than generally appreciated. In turn, this area deserves more attention than it currently receives. Mechanisms for interfibrillar stress transfer almost certainly involve the surface and interfibrillar proteoglycans, where molecular entanglement and electrostatic interactions may provide the basis for shear resistivity. The glycosaminoglycans bound to decorin act like bridges between contiguous fibrils connecting adjacent fibrils every 64–68 nm. This architecture would suggest their possible role in providing the mechanical integrity of the tendon structure. Such interactions have been investigated (among others) by Redaelli *et al.* (2003).

XII. Conclusions

The collagen fibril is a complex structure with a fundamental D-repeat of approximately 65 to 67 nm. This provides a framework on how different molecular species interact with one another, to provide fibrillar structures with features suited to both their roles as biomechanical tensile materials and as a source of intermolecular connectivity. The ubiquity of collagen in

most animal tissues shows that a high degree of functionality can be married by modulations on a relatively stable framework. The hierarchical organization of collagen-based materials from the molecular to the functional tissue allows for the intervention and interplay of a series of design features that ensure the ''triple-helical rope'' is put to best use in each case.

Future studies will focus on understanding the contribution of different hierarchical levels to the overall performance of the tissue. This in turn will allow the optimum design of biomimetics and future biomaterials, and will provide essential information for tissue engineering.

References

Adachi, E., and Hayashi, T. (1986). *In vitro* formation of hybrid fibrils of type-V collagen and type-I collagen—Limited growth of type-I collagen into thick fibrils by type-V collagen. *Connect. Tissue Res.* **14,** 257–266.

Bailey, A. J. (2001). Molecular mechanisms of ageing in connective tissues. *Mech. Ageing Dev.* **122,** 735–755.

Baselt, D. R., Revel, J. P., and Baldeschwieler, J. D. (1993). Subfibrillar structure of type-I collagen observed by atomic-force microscopy. *Biophys. J.* **65,** 2644–2655.

Birk, D. E. (2001). Type V collagen: Heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron* **32,** 223–237.

Birk, D. E., Zychband, E. I., Woodruff, S., Winkelmann, D. A., and Trelastd, R. L. (1997). Collagen fibrillogenesis *in situ*: Fibril segments become long fibrils as the developing tendon matures. *Dev. Dynam.* **208,** 291–298.

Bishop, P. N. (2000). Structural macromolecules and supramolecular organisation of the vitreous gel. *Prog. Retin. Eye Res.* **19,** 323–344.

Blaschke, U. K., Eikenberry, E. F., Hulmes, D. J. S., Galla, H. J., and Bruckner, P. (2000). Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. *J. Biol. Chem.* **275,** 10370–10378.

Bradshaw, J. P., Miller, A., and Wess, T. J. (1989). Phasing the meridional diffraction pattern of type I collagen using isomorphous derivatives. *J. Mol. Biol.* **205,** 685–694.

Bradshaw, A. D., Puolakkainen, P., Dasgupta, J., Davidson, J. M., Wight, T. N., and Sage, E. H. (2003). SPARC-null mice display abnormalities in the dermis characterized by decreased collagen fibril diameter and reduced tensile strength. *J. Invest. Dermatol.* **120,** 949–955.

Brodsky, B., Eikenberry, E. F., and Cassidy, K. (1980). An unusual collagen periodicity in skins. *Biochim. Biophys. Acta* **621,** 162–166.

Brown, E. M., King, G., and Chen, J. M. (1997). Model of the helical portion of a type I collagen microfibril. *J. Am. Leather Chem. As.* **92,** 1–7.

Cameron, G. J., Alberts, I. L., Laing, J. H., and Wess, T. J. (2002). Structure of type I and type III heterotypic collagen fibrils: An X-ray diffraction study. *J. Struct. Biol.* **137,** 15–22.

Chakravarti, S., Magnuson, T., Lass, J. H., Jepsen, K. J., LaMantia, C., and Carroll, H. (1998). Lumican regulates collagen fibril assembly: Skin fragility and corneal opacity in the absence of lumican. *J. Cell Biol.* **141,** 1277–1286.

Chanut-Delalande, H., Fichard, A., Bernocco, S., Garrone, R., Hulmes, D. J. S., and Ruggiero, F. (2001). Control of heterotypic fibril formation by collagen V is determined by chain stoichiometry. *J. Biol. Chem.* **276,** 24352–24359.

Chapman, J. A. (1989). The regulation of size and form in the assembly of collagen fibrils *in vivo*. *Biopolymers* **28,** 1367–1382.

Craig, A. S., Birtles, M. J., Conway, J. F., and Parry, D. A. D. (1989). An estimate of the mean length of collagen fibrils in rat tail tendon as a function of age. *Connect. Tissue Res.* **19,** 51–62.

Danielson, K. G., Siracusa, L. D., Donovan, P. J., and Iozzo, R. V. (1999). Decorin, epiphycan, and lumican genes are closely linked on murine chromosome 10 and are deleted in lethal steel mutants. *Mamm. Genome.* **10,** 201–203.

Derwin, K. A., Soslowsky, L. J., Kimura, J. H., and Plaas, A. H. (2001). Proteoglycans and glycosaminoglycan fine structure in the mouse tail tendon fascicle. *J. Orthop. Res.* **19,** 269–277.

Dreisewerd, K., Rohlfing, A., Spottke, B., Urbanke, C., and Henkel, W. (2004). Characterization of whole fibril-forming collagen proteins of types I, III, and V from foetal calf skin by infrared matrix-assisted laser desorption ionization mass spectrometry. *Anal. Chem.* **76,** 3482–3491.

Doyle, B. B., Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A., and Woodhead-Galloway, J. (1974). A D-periodic narrow filament in collagen. *Proc. R. Soc. Lond. B.* **186,** 67–74.

Eikenberry, E. F., Childs, B., Sheren, S. B., Parry, D. A. D., Craig, A. S., and Brodsky, B. (1984). Crystalline fibril structure of type-II collagen in lamprey notochord sheath. *J. Mol. Biol.* **176,** 261–277.

Eyre, D. R., Pietka, T., Weis, M. A., and Wu, J. J. (2004). Covalent cross-linking of the NC1 domain of collagen type IX to collagen type II in cartilage. *J. Biol. Chem.* **279,** 2568–2574.

Folkhard, W., Christmann, D., Geercken, W., Knorzer, E., Koch, M. H., Mosler, E., Nemetschek-Gansler, H., and Nemetschek, T. (1987a). Twisted fibrils are a structural principle in the assembly of interstitial collagens, chordae tendinae included. *Z. Naturforsch.* [C] **42,** 1303–1306.

Folkhard, W., Mosler, E., Geercken, E., Knorzer, H., Nemetschek-Gansler, H., and Nemetschek, T. (1987b). Quantitative analysis of the molecular sliding mechanism in native tendon collagen—Time resolved dynamic studies using synchrotron radiation. *Int. J. Biol. Macromol.* **9,** 169–175.

Franc, S. (1993). Ultrastructural evidences of a distinct axial domain within native rat tail tendon collagen fibrils. *J. Submicrosc. Cytol. Path.* **25,** 85–91.

Fraser, R. D. B., MacRae, T. P., Miller, A., and Suzuki, E. (1983). Molecular conformation and packing in collagen fibrils. *J. Mol. Biol.* **167,** 497–521.

Fraser, R. D. B., MacRae, T. P., and Miller, A. (1987). Molecular packing in type I collagen fibrils. *J. Mol. Biol.* **193,** 115–125.

Fratzl, P. (2003). Cellulose and collagen: from fibres to tissues. *Curr. Opin. Colloid In.* **8,** 32–39.

Fratzl, P., Fratzl-Zelman, N., and Klaushofer, K. (1993). Collagen packing and mineralization. An X-ray scattering investigation of turkey leg tendon. *Biophys. J.* **64,** 260–266.

Fratzl, P., Misof, K., Zizak, I., Rapp, G., Amenitsch, H., and Bernstorff, S. (1998). Fibrillar structure and mechanical properties of collagen. *J. Struct. Biol.* **122,** 119–122.

Giraud-Guille, M.-M. (1996). Twisted liquid crystalline supramolecular arrangements in morphogenesis. *Int. Rev. Cytol.* **166,** 59–101.

Giraud-Guille, M. M., Besseau, L., and Martin, R. (2003). Liquid crystalline assemblies of collagen in bone and *in vitro* systems. *J. Biomech.* **36,** 1571–1579.

Gorham, S. D., Light, N. D., Diamond, A. M., Willins, M. J., Bailey, A. J., Wess, T. J., and Leslie, N. J. (1991). Effect of chemical modifications on the susceptibility of collagen to proteolysis II. Dehydrothermal crosslinking. *Int. J. Biol. Macromol.* **14,** 129–138.

Gutsmann, T., Fantner, G. E., Venturoni, M., Ekani-Nkodo, A., Thompson, J. B., Kindt, J. H., Morse, D. E., Fygenson, D. K., and Hansma, P. K. (2003). Evidence that collagen fibrils in tendons are inhomogeneously structured in a tubelike manner. *Biophys. J.* **84,** 2593–2598.

Hagg, R., Bruckner, P., and Hedbom, E. (1998). Cartilage fibrils of mammals are biochemically heterogeneous: Differential distribution of decorin and collagen IX. *J. Cell Biol.* **142,** 285–294.

Hansen, U., and Bruckner, P. (2003). Macromolecular specificity of collagen fibrillogenesis. Fibrils of collagens I and XI contain a heterotypic alloyed core and a collagen I sheath. *J. Biol. Chem.* **278,** 37352–37359.

Haston, J. L., FitzGerald, O., Kane, D., and Smith, K. D. (2002). Preliminary observations on the influence of rheumatoid alpha-1–acid glycoprotein on collagen fibril formation. *Biomed. Chromatogr.* **16,** 332–342.

Hedlund, H., Mengarelli-Widholm, S., Heinegard, D., Reinholt, F. P., and Svensson, O. (1994). Fibromodulin distribution and association with collagen. *Matrix Biol.* **14,** 227–232.

Hedlund, H., Hedbom, E., Heinegard, D., Mengarelli-Widholm, S., Reinholt, F. P., and Svensson, O. (1999). Association of the aggrecan keratan sulfate-rich region with collagen in bovine articular cartilage. *J. Biol. Chem.* **274,** 5777–5781.

Helseth, D. L., and Veis, A. (1981). Collagen self-assembly *in vitro.* Differentiating specific telopeptide dependent interactions using selective enzyme modification and the addition of free amino telopeptide. *J. Biol. Chem.* **256,** 7118–7128.

Helseth, D. L., Jr., Lechner, J. H., and Veis, A. (1979). Role of the amino-terminal extrahelical region of type I collagen in directing the 4D overlap in fibrillogenesis. *Biopolymers* **18,** 3005–3014.

Henkel, W. (1996). Cross-link analysis of the C-telopeptide domain from type III collagen. *Biochem. J.* **318,** 497–503.

Henkel, W., and Glanville, R. W. (1982). Covalent crosslinking between molecules of type I and type III collagen. The involvement of the N-terminal, nonhelical regions of the alpha 1 (I) and alpha 1 (III) chains in the formation of intermolecular crosslinks. *Eur. J. Biochem.* **122,** 205–213.

Highberger, J. H., Gross, J., and Schmitt, F. O. (1950). Electron microscope observations of certain fibrous structures obtained from connective tissue extracts. *J. Am. Chem. Soc.* **72,** 3321–3322.

Hodge, A. J., and Petruska, J. A. (1963). Recent studies with the electron microscope on ordered aggregates of the tropocollagen molecule. *In* ''Aspects of Protein Chemistry'' (G. N. Ramachandran, Ed.), pp. 289–300. Academic Press, London.

Hofmann, H., Fietzek, P. P., and Kuhn, K. (1980). Comparative analysis of the sequences of the three collagen chains $\alpha 1$(I), $\alpha 2$ and $\alpha 1$(III); Function and genetic aspects. *J. Mol. Biol.* **141,** 293–314.

Holmes, D. F., Gilpin, C. J., Baldock, C., Ziese, U., Koster, A. J., and Kadler, K. E. (2001). Corneal collagen fibril structure in three dimensions: Structural insights

into fibril assembly, mechanical properties, and tissue organization. *Proc. Natl. Acad. Sci. USA* **98,** 7307–7312.

Hukins, D. W. L., and Woodhead-Galloway, J. (1977). Collagen fibrils as examples of smectic A biological fibres. *Mol. Cryst. Liq. Cryst.* **41,** 33–39.

Hulmes, D. J. S. (2002). Building collagen molecules, fibrils, and suprafibrillar structures. *J. Struct. Biol.* **137,** 2–10.

Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A., and Woodhead-Galloway, J. (1973). Analysis of the primary structure of collagen for the origins of molecular packing. *J. Mol. Biol.* **79,** 137–148.

Hulmes, D. J. S., Miller, A., White, S. W., and Brodsky-Doyle, B. (1977). Interpretation of the meridional diffraction pattern from collagen fibres in terms of the known amino acid sequence. *J. Mol. Biol.* **110,** 643–666.

Hulmes, D. J. S., Miller, A., White, S. W., Timmins, P. A., and Berthet-Colominas, C. (1980). Interpretation of the low angle meridional neutron diffraction patterns from collagen fibres in terms of the amino acid sequence. *Int. J. Biol. Macromol.* **2,** 338–345.

Hulmes, D. J. S., Holmes, D. F., and Cummings, C. (1985). Crystalline regions in collagen fibrils. *J. Mol. Biol.* **184,** 473–477.

Hulmes, D. J. S., Wess, T. J., Prockop, D. J., and Fratzl, P. (1995). Radial packing, order and disorder in collagen fibrils. *Biophys. J.* **68,** 1661–1670.

Hwang, W. S., Li, B., Jin, L. H., Ngo, K., Schachar, N. S., and Hughes, G. N. F. (1992). Collagen fibril structure of normal, aging, and osteoarthritic cartilage. *J. Pathol.* **167,** 425–433.

Itoh, T., Kobayashi, M., and Hashimoto, M. (1998). The role of intermolecular electrostatic interaction on appearance of the periodic band structure in type I collagen fibril. *Jpn. J. Appl. Phys. 1.* **37,** L190–L192.

Jesoir, J. C., Miller, A., and Berthet-Colominas, C. (1981). Crystalline three dimensional packing is a general feature of type I collagen fibrils. *FEBS Lett.* **13,** 238–240.

Jones, E. Y., and Miller, A. (1987). Models for the N-terminal and C-terminal telopeptide regions of interstitial collagens. *Biopolymers* **26,** 463–480.

Jones, E. Y., and Miller, A. (1991). Analysis of structural design features in collagen. *J. Mol. Biol.* **218,** 209–219.

Kamiyama, R. (1982). Fibrous long spacing-like fibers in the bone marrow of myeloproliferative disorder. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **39,** 285–291.

Kannus, P. (2000). Structure of the tendon connective tissue. *Scand. J. Med. Sci. Spor.* **10,** 312–320.

Kassner, A., Tiedemann, K., Notbohm, M., Ludwig, T., Morgelin, M., Reinhardt, D. P., Chu, M. L., Bruckner, P., and Grassel, S. (2004). Molecular structure and interaction of recombinant human type XVI collagen. *J. Mol. Biol.* **339,** 835–853.

Knight, D. P., and Vollrath, F. (2002). Biological liquid crystal elastomers. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.* **357,** 155–163.

Koch, M., Foley, J. E., Hahn, R., Zhou, P. H., Burgeson, R. E., Gerecke, D. R., and Gordon, M. K. (2001). Alpha 1(XX) collagen, a new member of the collagen subfamily, fibril-associated collagens with interrupted triple helices. *J. Biol. Chem.* **276,** 23120–23126.

Kuwaba, K., Kobayashi, M., Nomura, Y., Irie, S., and Koyama, Y. (2002). Size control of decorin dermatan sulfate during remodeling of collagen fibrils in healing skin. *J. Dermatol. Sci.* **29,** 185–194.

Landis, W. J., Hodgens, K. J., Song, M. J., Arena, J., Kiyonaga, S., Marko, M., Owen, C., and McEwen, B. F. (1996). Mineralization of collagen may occur on fibril surfaces:

Evidence from conventional and high-voltage electron microscopy and three-dimensional imaging. *J. Struct. Biol.* **117,** 24–35.

Linsenmayer, T. F., Gibney, E., Igoe, F., Gordon, M. K., Fitch, J. M., Fessler, L. I., and Birk, D. E. (1993). Type-V collagen: Molecular-structure and fibrillar organization of the chicken alpha-1(V) NH_2-terminal domain, a putative regulator of corneal fibrillogenesis. *J. Cell Biol.* **121,** 1181–1189.

Liu, X. H., Otter, A., Scott, P. G., Cann, J. R., and Kotovych, G. (1993). Conformational analysis of the type-II and type-III collagen alpha-1 chain C-telopeptides by H-1–nmr and Circular-Dichroism Spectroscopy. *J. Biomol. Struct. Dyn.* **11,** 541–555.

Lucic, D., Mollenhauer, J., Kilpatrick, K. E., and Cole, A. A. (2003). N-telopeptide of type II collagen interacts with annexin V on human chondrocytes. *Connect. Tissue Res.* **44,** 225–239.

MacBeath, J. R., Shackleton, D. R., and Hulmes, D. J. S. (1993). Tyrosine-rich acidic matrix protein (TRAMP) accelerates collagen fibril formation *in vitro. J. Biol. Chem.* **268,** 19826–19832.

McBride, D. J., Choe, V., Shapiro, J. R., and Brodsky, B. (1997). Altered collagen structure in mouse tail tendon lacking the alpha 2(I) chain. *J. Mol. Biol.* **270,** 275–284.

Malone, J. P., George, A., and Veis, A. (2004). Type I collagen N-telopeptides adopt an ordered structure when docked to their helix receptor during fibrillogenesis. *Proteins* **54,** 206–215.

Marchini, M., Morocutti, M., Ruggeri, A., Koch, M. H. J., Bigi, A., and Roveri, N. (1986). Differences in the fibril structure of corneal and tendon collagen. An electron microscopy and X-ray diffraction investigation. *Connect. Tissue Res.* **15,** 269–281.

Martin, R., Farjanel, J., Eichenberger, D., Colige, A., Kessler, E., Hulmes, D. J. S., and Giraud-Guille, M. M. (2000). Liquid crystalline ordering of procollagen as a determinant of three-dimensional extracellular matrix architecture. *J. Mol. Biol.* **301,** 11–17.

Meek, K. M., and Fullwood, N. J. (2001). Corneal and scleral collagens—A microscopist's perspective. *Micron* **32,** 261–272.

Meek, K. M., Chapman, J. A., and Hardcastle, R. A. (1979). The staining pattern of collagen fibrils. Improved correlation with sequence data. *J. Biol. Chem.* **254,** 10710–10714.

Michna, H. (1984). Morphometric analysis of loading-induced changes in collagen-fibril populations in young tendons. *Cell Tissue Res.* **236,** 465–470.

Miller, A., and Tochetti, D. (1981). Calculated X-ray diffraction pattern from a quasi-hexagonal model for the molecular arrangement in collagen. *Int. J. Macromol.* **3,** 9–18.

Mizuno, K., Adachi, E., Imamura, Y., Katsumata, O., and Hayashi, T. (2001). The fibril structure of type V collagen triple-helical domain. *Micron* **32,** 317–323.

Morris, C. J., Bradby, G. V., and Walton, K. W. (1978). Fibrous long-spacing collagen in human atherosclerosis. *Atherosclerosis* **31,** 345–354.

Nakanishi, S., Masuda, T., Kitamura, T., Moriizumi, K., and Kajikawa, T. (1981). Distribution of fibrous long-spacing fibers in normal and pathological lymph nodes. *Acta Pathol. Jpn.* **31,** 733–745.

North, A. C. T., Cowan, P. M., and Randall, J. T. (1954). Structural units in collagen fibrils. *Nature* **174,** 1142–1143.

Oguma, H., Murakami, G., Takahashi-Iwanaga, H., Aoki, M., and Ishii, S. (2001). Early anchoring collagen fibres at the bone-tendon interface are conducted by woven bone formation. *J. Orthopaedic Research* **19,** 873–880.

Olsen, B. R. (1997). Collagen IX. *Int. J. Biochem. Cell Biol.* **29,** 555–558.

Orgel, J. P., Wess, T. J., and Miller, A. (2000). The *in situ* conformation and axial location of the intermolecular cross-linked non-helical telopeptides of type I collagen. *Struct. Fold. Des.* **8,** 137–142.

Orgel, J. P. R. O., Miller, A., Irving, T. C., Fischetti, R. F., Hammersley, A. P., and Wess, T. J. (2001). The *in situ* supermolecular structure of type I collagen. *Structure* **9,** 1061–1069.

Ortolani, F., Giordano, M., and Marchini, M. (2000). A model for type II collagen fibrils: Distinctive D-band patterns in native and reconstituted fibrils compared with sequence data for helix and telopeptide domains. *Biopolymers* **54,** 448–463.

Ottani, V., Raspanti, M., and Ruggeri, A. (2001). Collagen structure and functional implications. *Micron* **32,** 251–260.

Otter, A., Scott, P. G., and Kotovych, G. (1988). Type-I collagen α-1chain C-telopeptide—Solution structure determined by 600 MHz proton NMR spectroscopy and implications for its role in collagen fibrillogenesis. *Biochemistry* **27,** 3560–3567.

Paige, M. F., Rainey, J. K., and Goh, M. C. (2001). A study of fibrous long spacing collagen ultrastructure and assembly by atomic force microscopy. *Micron* **32,** 341–353.

Parkinson, J., Kadler, K. E., and Brass, A. (1995). Simple physical model of collagen fibrillogenesis based on diffusion-limited aggregation. *J. Mol. Biol.* **247,** 823–831.

Parry, D. A. D., Barnes, G. R. G., and Craig, A. S. (1978). A comparison of the size distribution of collagen fibrils in connective tissues as a function of age and a possible relationship between fibril size distribution and mechanical properties. *Proc. Roy. Soc. Lond. B.* **203,** 305–321.

Parry, D. A. D., and Craig, A. S. (1979). Electron microscope evidence for an 80 Angstrom unit in collagen fibrils. *Nature* **282,** 213–225.

Parry, D. A. D., and Craig, A. S. (1984). Growth and Development of Collagen Fibrils in Connective Tissues. *In* ''Ultrastructure of the Connective Tissue Matrix'' (A. Ruggeri and P.M. Motta, Eds.), pp. 34–64. Martinus Nijhoff, Netherlands.

Paterlini, M. G., Nemethy, G., and Scheraga, H. A. (1995). The energy of formation of internal loops in triple-helical collagen polypeptides. *Biopolymers* **35,** 607–619.

Patterson-Kane, J. C., Wilson, A. M., Firth, E. C., Parry, D. A. D., and Goodship, A. E. (1997). Comparison of collagen fibril populations in the superficial digital flexor tendons of exercised and non-exercised thoroughbreds. *Equine Vet. J.* **29,** 121–125.

Pihlajamaa, T., Lankinen, H., Ylostalo, J., Valmu, L., Jaalinoja, J., Zaucke, F., Spitznagel, L., Gosling, S., Puustinen, A., Morgelin, M., Peranen, J., Maurer, P., Ala-Kokko, L., and Kilpelainen, I. (2004). Characterization of recombinant amino-terminal NC4 domain of human collagen IX: Interaction with glycosaminoglycans and cartilage oligomeric matrix protein. *J. Biol. Chem.* **279,** 24265–24273.

Pinheiro, M. C., Moraes, S. G., Battlehner, C. N., Caldini, E. G., Toledo, O. M. S., and Joazeiro, P. P. (2004). Histochemical and ultrastructural study of collagen fibers in mouse pubic symphysis during late pregnancy. *Micron.* **35,** 685–693.

Porter, K. R., and Pappas, G. D. (1958). Collagen formation by fibroblasts of the chick embryo dermis. *J. Biophysic. Biochem. Cytol.* **5,** 153–180.

Prockop, D. J., and Hulmes, D. J. S. (1994). Assembly of collagen fibrils *de novo* from soluble precursors. *In* ''Extracellular Matrix Assembly and Structure'' (P. D. Yurchenco, D. E. Birk and R. P. Mecham, Eds.), pp. 47–90. Academic Press, San Diego.

Raspanti, M., Alessandrini, A., Gobbi, P., and Ruggeri, A. (1996). Collagen fibril surface: TMAFM, FEG-SEM and freeze-etching observations. *Microscopy Res. Tech.* **35,** 87–93.

Redaelli, A., Vesentini, S., Soncini, M., Vena, P., Mantero, S., and Montevecchi, F. M. (2003). Possible role of decorin glycosaminoglycans in fibril to fibril force transfer in relative mature tendons: A computational study from molecular to microstructural level. *J. Biomech.* **36,** 1555–1569.

Ronziere, M-C., Berthet–Colominas, C., and Herbage, D. (1987). Comparative structural studies of reconstituted and native type I and type II collagen fibrils by low angle X-ray diffraction. *Biochim. Biophys. Acta* **916,** 381–387.

Ronziere, M-C., Herbage, B., Herbage, D., and Bernengo, J-C. (1998). Fourier analysis of electron micrographs of positively stained collagen fibrils: Application to type I and II collagen typing. *Int. J. Biol. Macromol.* **23,** 207–213.

Sanders, J. E., and Goldstein, B. S. (2001). Collagen fibril diameters increase and fibril densities decrease in skin subjected to repetitive compressive and shear stresses. *J. Biomech.* **34,** 1581–1587.

Sasaki, N., and Odajima, S. (1996). Elongation mechanism of collagen fibrils and force-strain relations of tendon at each level of structural hierarchy. *J. Biomech.* **29,** 1131–1136.

Sasaki, N., Shukunami, N., Matsushima, N., and Izumi, Y. (1999). Time resolved X-ray diffraction from tendon collagen during creep using synchrotron radiation. *J. Biomech.* **32,** 285–292.

Scott, J. E., and Thomlinson, A. M. (1998). The structure of interfibrillar proteoglycan bridges ('shape modules') in extracellular matrix of fibrous connective tissues and their stability in various chemical environments. *J. Anat.* **192,** 391–405.

Silver, F. H., Christiansen, D. L., Snowhill, P. B., and Chen, Y. (2001). Transition from viscous to elastic-based dependency of mechanical properties of self-assembled type I collagen fibers. *J. Appl. Polymer Sci.* **79,** 134–142.

Slavin, R. E., Swedo, J. L., Brandes, D., Gonzalez-Vitale, G. C., and Osornio-Vargas, A. (1985). Extrapulmonary silicosis: A clinical, morphologic, and ultrastructural study. *Hum. Pathol.* **16,** 393–412.

Smith, J. W. (1968). Molecular pattern in native collagen. *Nature* **219,** 157–158.

Svensson, L., Aszodi, A., Reinholt, F. P., Fassler, R., Heinegard, D., and Oldberg, A. (1999). Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered Lumican deposition in tendon. *J. Biol. Chem.* **274,** 9636–9647.

Trus, B. L., and Piez, K. A. (1980). Compressed microfibril models of the native collagen fibril. *Nature* **286,** 300–301.

Vaughan, L., Mendler, M., Huber, S., Bruckner, P., Winterhalter, K. H., Irwin, M. I., and Mayne, R. (1988). D-periodic Distribution of collagen type IX along cartilage fibrils. *J. Cell Biol.* **106,** 991–997.

Venturoni, M., Gutsmann, T., Fantner, G. E., Kindt, J. H., and Hansma, P. K. (2003). Investigations into the polymorphism of rat tail tendon fibrils using atomic force microscopy. *Biochem. Biophys. Res. Comm.* **303,** 508–513.

Vesentini, S., Redaelli, A., and Montevecchi, F. M. (2005). Estimation of the binding force of the collagen molecule-decorin core protein complex in collagen fibril. *J. Biomech.* **38,** 433–443.

Vitagliano, L., Nemethy, G., Zagari, A., and Scheraga, H. A. (1995). Structure of the type-I collagen molecule based on conformational energy computations: The triple-stranded helix and the N-terminal telopeptide. *J. Mol. Biol.* **247,** 69–80.

Watanabe, M., Kobayashi, M., Fujita, Y., Senga, K., Mizutani, H., Ueda, M., and Hoshino, T. (1997). Association of type VI collagen with D-periodic collagen fibrils in developing tail tendons of mice. *Arch. Histol. Cytol.* **60,** 427–434.

Wen, C. K., and Goh, M. C. (2004). AFM nanodissection reveals internal structural details of single collagen fibrils. *Nano. Lett.* **4,** 129–132.

Wess, T. J., Hammersley, A., Wess, L., and Miller, A. (1995). Type I collagen packing conformation of the triclinic unit cell. *J. Mol. Biol.* **248,** 487–493.

Wess, T. J., Hammersley, A. P., Wess, L., and Miller, A. (1998). A consensus model for molecular packing of type I collagen. *J. Struct. Biol.* **122,** 92–100.

Wessel, H., Anderson, S., Fite, D., Halvas, E., Hempel, J., and SundarRaj, N. (1997). Type XII collagen contributes to diversities in human corneal and limbal extracellular matrices. *Invest. Ophthalmol. Vis. Sci.* **38,** 2408–2422.

White, J., Werkmeister, J. A., Ramshaw, J. A. M., and Birk, D. E. (1997). Organization of fibrillar collagen in the human and bovine cornea: Collagen types V and III. *Connect. Tissue Res.* **36,** 165–174.

Yamamoto, T., and Wakita, M. (1992). Bundle formation of principal fibers in rat molars. *J. Periodont. Res.* **27,** 20–27.

Yamamoto, S., Hashizume, H., Hitomi, J., Shigeno, M., Sawaguchi, S., Abe, H., and Ushiki, T. (2000a). The subfibrillar arrangement of corneal and scleral collagen fibrils as revealed by scanning electron and atomic force microscopy. *Arch. Histol. Cytol.* **63,** 127–135.

Yamamoto, T., Domon, T., Takahashi, S., Islam, N., and Suzuki, R. (2000b). Twisted plywood structure of an alternating lamellar pattern in cellular cementum of human teeth. *Anat. Embryol.* **202,** 25–30.

Young, R. D., Lawrence, P. A., Duance, V. C., Aigner, T., and Monaghan, P. (2000a). Immunolocalization of collagen types II and III in single fibrils of human articular cartilage. *J. Histochem. Cytochem.* **48,** 423–432.

Young, B. B., Gordon, M. K., and Birk, D. E. (2000b). Expression of type XIV collagen in developing chicken tendons: Association with assembly and growth of collagen fibrils. *Dev. Dyn.* **217,** 430–439.

MOLECULAR PACKING IN NETWORK-FORMING COLLAGENS

By CARLO KNUPP* AND JOHN M. SQUIRE†

*Structural Biophysics Group, School of Optometry and Vision Sciences, Cardiff University, Cardiff CF10 3NX, United Kingdom;
†Biological Structure and Function Section, Biomedical Sciences Division, Imperial College Faculty of Medicine, London SW7 2AZ, United Kingdom

I. Introduction 375
II. Network-Forming Collagens in Detail 377
A. Type IV Collagen 377
B. Type VI Collagen 380
C. Type VIII Collagen 386
D. Type X Collagen 387
E. Dogfish Egg Case Collagen 388
III. Intertwining of Network Collagens: The Type VI Segmented Supercoil 392
IV. Common Structural Themes in Network-Forming Collagens 396
References 399

Abstract

Different collagen types can vary considerably in length, molecular weight, chemical composition, and the way they interact with each other to form molecular aggregates. Collagen Types IV, VI, VIII, X, and dogfish egg case collagen make linear and lateral associations to form open networks rather than fibers. The roles played by these network-forming collagens are diverse: they can act as support and anchorage for cells and tissues, serve as molecular filters, and even provide protective permeable barriers for developing embryos. Their functional properties are intimately linked to their molecular organization. This Chapter reviews what is known about the molecular structure of this group of collagens, describes the ways the molecules interact to form networks, and—despite the large variations in molecular size—identifies common aggregation themes.

I. Introduction

There are more than 25 distinct types of collagen; their amino acid sequences and molecular weights can vary considerably from one type to another. Not surprisingly, then, individual types of collagen can aggregate

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70011-X

Copyright 2005, Elsevier Inc.
All rights reserved.

0065-3233/05 $35.00

to form different types of structures. Types IV, VI, VIII, X, and dogfish egg case collagen, for example, can aggregate linearly and laterally to give rise to open networks. These collagens are present in smaller proportions than Type I collagen, but their importance is underlined by the variety of essential roles that they play in organisms. In fact, collagen networks act as supporting structures for cells and tissues, serve as selective molecular filters and barriers, function as anchorage for neighboring cells, and (in the case of the networks in dogfish egg cases) contain and protect developing embryos.

Of the network-forming collagens, collagen Type IV is arguably the most important as the main component of basement membranes (see below). It is a relatively long molecule (about 400 nm long) that assembles irregularly when forming networks. Since structural biologists need regularity in the systems that they investigate, from a structural point of view collagen IV is very difficult to analyze. Nonetheless, by studying limited regions of the molecule (e.g., the N- and C-terminal regions) and the general modes of lateral interaction, it has been possible to arrive at a working model for the construction of collagen IV networks.

Collagens Type VIII and X form hexagonal networks. They are shorter than collagen IV and the networks that they form are more regular. However, no detailed three-dimensional (3D) description exists of the networks that they form in terms of crystallographic symmetry and stoichiometry.

For a long time, Type VI collagen was not generally considered as part of the network-forming collagen family. In the past, collagen VI tetramers were described to aggregate linearly to form so-called *beaded filaments*. It now appears that, as well as aggregating linearly, they are also able to associate laterally through their globular domains, thus forming 3D networks. It is not certain whether these lateral associations can occur spontaneously, or if they are mediated by other molecules. However, unlike collagens IV and VIII in basement membranes, collagen VI lateral networks may not necessarily have physiologically important structural roles. It is possible that they arise as a consequence of aging or pathological conditions. Nevertheless, their analysis can throw light on the general structural patterns of collagen network formation, as well as improve our knowledge of the pathological or aging processes with which they are associated. Type VI filaments, on the other hand, play important physiological roles; Type VI collagen mutations have been shown to be linked to Bethlem myopathy (Scacheri *et al.*, 2002) and Ullrich syndrome (Camacho Vanegas *et al.*, 2001).

Dogfish egg case collagen is relatively short (about 45 nm) and assembles in a remarkably regular network. It has proved a good model for the study of collagen network formation, in particular collagen VI networks.

Nematocyst minicollagens have also been reported. These are the shortest known collagens, with only 14 Gly-X-Y repeats. They are the main constituents of the capsule wall of the nematocyst, an explosive organelle used by cnidaria such as hydras, jellyfish, and corals for defensive, aggressive, or locomotive purposes (Engel *et al.,* 2001; Holstein *et al.,* 1994). Although exhaustive structural work has not yet been carried out, it is possible that these collagens form networks by crosslinking linearly through cysteine-rich globular domains (Ozbek *et al.,* 2002; Pokidysheva *et al.,* 2004; Skaer and Picken, 1965).

This Chapter contains a description of the main network-forming collagens. It highlights the main structural characteristics of each individual collagen that allow network formation, and ends by discussing the underlying structural commonalities between different network collagen types.

II. Network-Forming Collagens in Detail

A. *Type IV Collagen*

Type IV collagen is a fundamental component of basement membranes (Fig. 1A). Basement membranes represent the portion of extracellular matrix that remains in direct contact with its formative cells. Basement membranes play an important role in cell adhesion, growth and differentiation, tissue repair, molecular ultrafiltration, cancer cell invasion, and metastasis.

Type IV collagen is composed of three 400-nm-long polypeptide chains. The triple-helical domains of the polypeptide chains are often interrupted (Fig. 1B). In humans, for example, the $\alpha 1$ chain has 21 interruptions of the Gly-X-Y triplet sequence, and in the $\alpha 2$ chain there are 23 interruptions. Overlaps between these Gly-X-Y discontinuities produce 26 irregularly spaced interruptions, resulting in a collagen that is flexible (Brazel *et al.,* 1988). The invariance in location of many of these interruptions among different species suggests a functional role (Timpl, 1989), but few functions have yet been directly demonstrated for the interruptions. From an amino acid sequence analysis of collagen VI (Knupp and Squire, 2001), it is apparent that interruptions of the Gly-X-Y repeat in this type of collagen reduce the stiffness of the molecule and help supercoiling. Gly-X-Y interruptions may play a similar role in collagen IV. It also appears likely that the flexibility generated by the interruptions is essential to allow the formation of a network rather than a fiber (Yurchenco and Ruben, 1987).

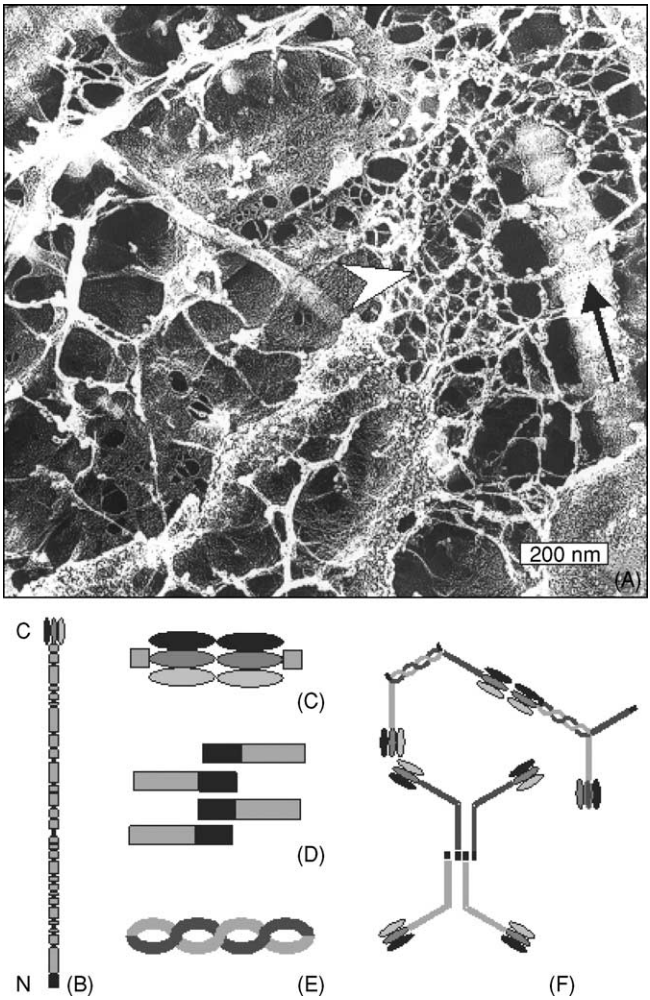


FIG. 1. (A) Metal-shadowed replica of partially purified glomerular basement membrane. The arrowhead points to a collagen IV network. The black arrow indicates a Type I collagen fiber. (Micrograph courtesy of Dr. M. Chew.) (B) Diagram of a collagen IV molecule. Along the triple-helical part (represented by gray rectangles), numerous interruptions of the Gly-X-Y amino acid repeat are present. Linear polymerization is observed through (C) C-terminal interactions and (D) N-terminal interactions. (E) Lateral interactions are also observed with apparent molecular supercoiling. (F) Collagen IV is able to form networks by linear and lateral associations.

Collagen IV $\alpha 1$ and $\alpha 2$ chains are found in all mammalian basement membranes (Borza *et al.*, 2001; Sado *et al.*, 1998). Variant Type IV collagen $\alpha 3$, $\alpha 4$, $\alpha 5$, and $\alpha 6$ chains have been identified (Butkowsky *et al.*, 1987; Oohashi *et al.*, 1994) and are present in a more restricted tissue distribution. The $\alpha 3$, $\alpha 4$, and $\alpha 5$ chains are found in the renal glomerular basement membrane and pulmonary alveolar basement membrane. The $\alpha 5$ and $\alpha 6$ chains are found in smooth muscle, aorta, bladder, mammalian duct and lobule, and epidermis basement membranes. The $\alpha 3$, $\alpha 4$, and $\alpha 5$ chains are thought to be responsible, when defective, for Alport syndrome (Sado *et al.*, 1998). There is evidence that collagen IV molecules in basement membranes exist mainly in three forms: $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$, $[\alpha 3(\mathrm{IV})][\alpha 4(\mathrm{IV})][\alpha 5(\mathrm{IV})]$, and $[\alpha 5(\mathrm{IV})]_2[\alpha 6(\mathrm{IV})]$ (Borza *et al.*, 2001; Sado *et al.*, 1998), although other combinations have been suggested (Kahsai *et al.*, 1997; Kalluri and Cosgroves, 2000).

Type IV collagen can assemble into a stable three-dimensional basement membrane network via three types of interactions. In the first type, pairs of C-terminal globules join together to form dimers (Fig. 1C). Studies of the $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$ form suggested that the dimeric globular domain is formed by disulphide exchange between corresponding cystines of the two monomeric domains (Siebold *et al.*, 1988). It was suggested that monomers would dimerize intracellularly through regulation of the redox conditions, which may represent the first step of higher-order assembly. This assembly step appeared to be driven by interactions of cysteines present in the globular sequences which form the dimers by disulphide exchanges in the $\alpha 1(\mathrm{IV})$ globular domains (Yurchenco and O'Rear, 1993). Recent results from crystallographic experiments did not confirm this hypothesis, but suggested that C-terminal interactions may be stabilized by a putative covalent Met-Lys crosslink (Than *et al.*, 2002), mediated by metal ions, or that they are the result of posttranslational modifications (Vanacore *et al.*, 2004).

It has been proposed that the three different forms of collagen IV arising from the combination of the six α-chains interact at the C-terminals in a restricted way: $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$ associates with another $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$ or with a $[\alpha 5(\mathrm{IV})]_2[\alpha 6(\mathrm{IV})]$; and $[\alpha 3(\mathrm{IV})][\alpha 4(\mathrm{IV})][\alpha 5(\mathrm{IV})]$ associates only with another $[\alpha 3(\mathrm{IV})][\alpha 4(\mathrm{IV})][\alpha 5(\mathrm{IV})]$ (Borza *et al.*, 2001; Boutaud *et al.*, 2000; Gunwar *et al.*, 2001; Sundaramoorthy *et al.*, 2002). The C-terminus would play a fundamental role in the assembly of chain-specific networks during both the initial alignment and selection of α-chains for molecular assembly, and during the selection and connection of molecules for network assembly (Sundaramoorthy *et al.*, 2002). Investigations of the $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$ form indicated that, in the second kind of interaction, four N-terminal ends bind to each other to form a 28-nm end-overlapped domain known as the *7S domain* (Fig. 1D). Assembly proceeds through

intermediates made of two or three antiparallel monomers held together through cooperative noncovalent interactions. A hydrophobic region on each N-terminal end has been identified from the primary sequence. The maximization of the hydrophobic contacts of the N-termini from four antiparallel molecules explains the antiparallel interactions with correct axial and azimuthal orientation. The complex is limited to tetramer size by the orientation arising from these interactions; cysteine and lysine/hydroxylysine residues are placed in the correct positions on the corresponding chains to form disulfide and nonreducible crosslinks (Siebold *et al.,* 1987). The above two interactions can be considered the basis for network formation.

In the third kind of interaction, Type IV collagen dimers self-interact through lateral (side-by-side) associations (Fig. 1E; Timpl and Brown, 1996; Yurchenco and Furthmayr, 1984; Yurchenco and Schittny, 1990). This interaction was first identified *in vitro* where assemblies showing extensive irregular networks were seen to form. The irregular polygonal geometry of the network suggests that collagen IV associates in more than one way, thus creating different spatial relationships between molecules. The existence of laterally associated networks (Fig. 1F) has been confirmed in tissue basement membranes of the human amnion (Timpl and Brown, 1996). In the electron microscope, these collagenous networks appeared as an extensive irregular polygonal network with three- to five-arm branch points, and with strands 2.5 to 7 nm in diameter. The average distance between branch points was about 45 ± 30 nm; integral globular domains, the same shape and size as those of purified collagens dimers, were also seen. Lateral associations were seen between single collagen IV filaments, with the formation of branching strands of variable diameter. Unfortunately, the localization of the 7S domains was difficult in both the reconstructed polymers and in basement membranes, possibly because of superimposed lateral associations. The presence of supramolecular helices of collagen IV filaments was seen in amniotic networks. The similarity between reconstituted and tissue basement membrane collagen networks suggests that the information for assembly is encoded in the collagen molecules themselves. The covalent bonding of mammalian collagenous networks is formed at the N- and C-terminal regions, but the loops formed by these ends would be expected to irreversibly entrap and stabilize helically wrapped, laterally associated filaments.

B. Type VI Collagen

Collagen VI is a rodlike molecule about 105-nm long with two globular regions at the extremities (Fig. 2A; Furthmayr *et al.,* 1983). The globular regions contain several domains homologous to von Willebrand

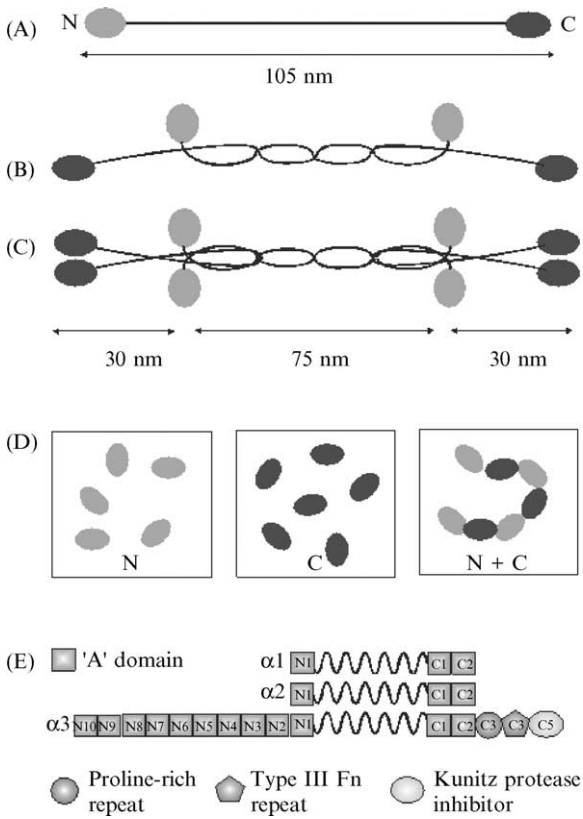


FIG. 2. (A) Diagram of a collagen VI molecule. Collagen VI is a 105 nm long molecule with massive globular domains at the extremities. (B) Two molecules can interact laterally with a 30 nm axial shift to form dimers. Intertwining of the monomers in the adjacent regions is seen. (C) Dimers join laterally to form tetramers. (D) Cleaved globular domains join linearly, N-termini with C-termini and vice versa, but homotypic N- or C-terminal associations are not seen. (E) Subdomain structure of the three Type VI collagen chain N- and C-terminal globular regions (after Baldock *et al.*, 2003).

factor A-domains (Fig. 2E; Baldock *et al.*, 2003). Collagen VI occurs mainly as a heterotrimer made of three different α-chains, but there is evidence for the existence of less stable assemblies comprising either $\alpha3$(VI) or $\alpha1$(VI)/$\alpha2$(VI) chains. All three α-chains have triple-helical sequences of similar length, but the $\alpha1$(VI) and $\alpha2$(VI) chains are relatively small, with only two *A* domains at each end. The $\alpha3$(VI) chain is much larger because of a prominent N-terminal region containing 10 *A* domains (Fig. 2E) and a C-terminus containing a proline-rich repeat, a fibronectin-like repeat,

and a Kunitz protease inhibitor (Baldock *et al.*, 2003). Type VI collagen microfibrils are stabilized by intra- and intermolecular disulphide bonds (Ayad *et al.*, 1994). Monomers are assembled intracellularly, first into antiparallel dimers and then into tetramers (Fig. 2B, C; Ayad *et al.*, 1994). Purified molecules are also seen to aggregate as antiparallel dimers. The monomers associate in an antiparallel fashion with a 30-nm axial shift. Dimers consist of an inner thick rodlike region 75-nm long, in which the two monomers are seen to intertwine four or five times, with two shorter (30 nm) and thinner rodlike segments emerging axially from the inner region (Fig. 2B; Furthmayr *et al.*, 1983). The sequences of the triple-helical parts of the Type VI collagen chains have been analyzed in detail, and it has been suggested (Knupp and Squire, 2001) that the three chains can twist into a segmented supercoil, details of which will be given below in Section III.

The globular ends of the two monomers are positioned at the extremities of the inner region and outer segments. Dimers can associate side-by-side with their 30-nm long outer segments to form tetramers (Fig. 2C; Furthmayr *et al.*, 1983; Von Der Mark *et al.*, 1984). The formation of dimers occurs also through the interaction of the metal ion-dependent adhesion site in the α2C2 A-domains in the C-terminal with a GER sequence found in the α2(VI) chain (Ball *et al.*, 2003). Experiments carried out on cleaved N- and C-globular domains showed that they join linearly, side-by-side, one type connected to the other (Fig. 2D; Kuo *et al.*, 1995). In addition, the tetramers are seen to interact to form long chains (beaded filaments). Dimers, tetramers, and polymeric chains all have 30-75-30 nm spacing between the globular ends (Furthmayr *et al.*, 1983; Von Der Mark *et al.*, 1984; Wu *et al.*, 1987). Collagen VI is found to bind to hyaluronan (Kielty *et al.*, 1992), biglycan and decorin (Wiberg *et al.*, 2001), fibrillin (Ueda and Yue, 2003), and other matrix constituents.

Because of the variety of ways in which collagen VI molecules can interact with each other and to several matrix constituents, they are good candidates to form networks. Lateral association of beaded filaments (likely to be collagen VI filaments) with their globular domains in register were reported as early as the 1960s (Luse, 1960) in association with brain tumors. More recently, more compact networks of collagen VI were seen in the eyes of people suffering from full-thickness macular holes (Fig. 3A, B; Knupp *et al.*, 2000; Reale *et al.*, 2001), from age-related macular degeneration (Fig. 3C, D; Knupp *et al.*, 2002a), and from Sorsby's fundus dystrophy (Fig. 3E, F; Knupp *et al.*, 2002b). Although these ocular structures can be easily interpreted in terms of collagen VI tetramers, they differ in detail depending on the particular illness with which they are associated. Fig. 3A and B illustrates a typical appearance of the collagen structure found in

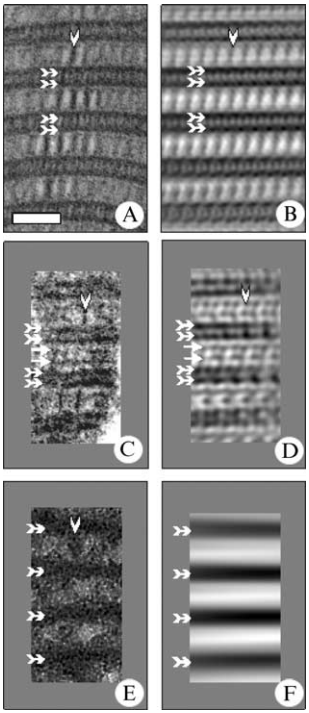


FIG. 3. (A) Collagen VI network in the vitreous of a patient suffering from full-thickness macular holes. Transverse bands 30 nm apart (double arrows) repeat axially with a 100 nm periodicity. Axial filaments run through the bands (chevron). (Bar represents 10 nm and applies to all parts of the figure). (B) Fourier-filtered image of (A). (C) Collagen VI network in the retina of a patient suffering from age-related macular degeneration. In addition to the structural characteristics of the aggregates described in (A), extra transverse bands of protein density run between the main pairs of bands (single arrows). (D) Fourier-filtered image of (C). (E) Collagen VI network in the retina of a patient suffering from Sorsby's fundus dystrophy. Here the transverse double bands are fused into a single broad band (double arrows). (F) Fourier-filtered image of (E). (Figure from Knupp and Squire, 2003.)

the vitreous of a patient suffering from full-thickness macular holes (Knupp *et al.*, 2000). Pairs of transverse bands, 30 nm apart (arrows), are repeated axially every ~100 nm. These bands arise from the globular domains at the extremities of the collagen VI tetramer outer segments. Filaments of protein density (chevron) run axially through the pairs of bands. In some views, their lateral periodicity is about 25 nm in the bigger gap between the transverse bands, and half that much in the smaller gap between the transverse bands. The filaments are staggered laterally by a

half-unit cell when traversing pairs of bands. Cross-sectional views of this structure showed dots of protein density arranged on a body-centered square array with 25 nm sides.

The model proposed for the collagen VI assembly in these structures (Fig. 4A) consists of collagen VI tetramers (or pairs of tetramers) interacting with four other tetramers (octamers) via their globular domains. These globular interactions involved C-termini interacting with N-termini and vice versa. In doing so, the outer segments of two collagen VI tetramers belonging to two different levels lie side by side, and give rise to the axial filaments seen in the smaller gap between pairs of transverse bands. In cross-section, the tetramer inner segments give rise to the axial filaments in the bigger gap between pairs of bands. Tetramers belonging to one level lie on a square lattice with 25 nm sides. Those belonging to the next level up (or down) also lie on a square lattice having the same dimensions, but are shifted laterally by a half-unit cell so that the tetramers are positioned in the center of the square lattice in the adjacent levels. This conformation accounts for both the lateral shift of the axial filaments in the bigger gaps between bands, and the body-centered square lattice seen in transverse views. A similar arrangement can be envisaged for the collagen VI assemblies associated with age-related macular degeneration (Fig. 3C, D).

In addition to the transverse bands of protein density seen in the aggregates described above, some of the assemblies associated with age-related macular degeneration present an extra pair of bands situated between and equidistant from the main pairs of bands (Fig. 3C, D, single arrows). The profiles of the main and subsidiary pairs of bands are the same, although the subsidiary pair is less electron dense. Models were developed to account for the subsidiary pairs of bands (Fig. 4B), and the preferred structure involved the presence of collagen VI dimers alongside the collagen VI tetramers. The dimers are required to be shifted axially by about 50 nm, and their globular domains would give rise to the secondary set of pairs in projection. It is likely that the dimer/tetramer association is mediated by retinal components. Extra material interacting with the tetramer outer segments, but not with those of the dimers, can account for the presence of additional mass that makes the main pairs of bands more electron dense.

In the collagen VI aggregates associated with Sorsby's fundus dystrophy, extra material—analogous to that described above, and tentatively identified as tissue inhibitor metalloproteinase 3 (TIMP3; abundant and defective in people suffering from this disease)—was proposed to bind to the outer segment of tetramers. It would thus participate in the further aggregation of collagen VI tetramers (Fig. 3E, F; Fig. 4C).

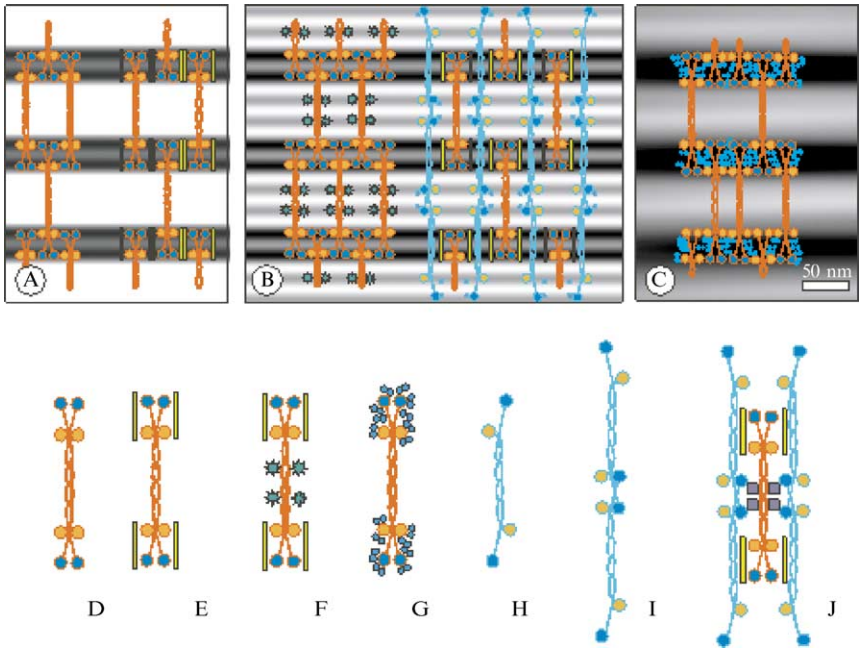


FIG. 4. Models for the collagen VI tetramer packing superimposed on one-dimensional averages of the electron densities from the structures described in Fig. 3A, C, and D, with a summary of the possible collagen VI structural interactions that form the building blocks that make networks. Collagen VI tetramers are represented in red. N- and C-termini are represented in blue and orange. Dimers are in light blue. (A) Left, model for the assembly associated with full-thickness macular holes; the packing occurs through the globular domain interactions, N-terminals interacting with the C-terminals and vice versa. Right, extra material (yellow bars) binding to the collagen VI tetramer outer segments could increase the electron density in this region. (B) Model for the assembly associated with age-related macular degeneration; the extra pairs of bands may arise from globular proteins binding to the tetramer triple helices (left) or more likely from the globular domains of dimers associating laterally with the tetramers (right). (C) Model for the assembly associated with Sorsby's fundus dystrophy. TIMP3 molecules (represented by blue circles) associate with the tetramer outer segments and fill the space between the bands. (D) Collagen VI tetramer. (E) As in (D), but with extra material binding to the tetramer outer segments (yellow bars). (F) As in (E), but with globular proteins that may account for the formation of extra bands of protein density. (G) As in (D), but with TIMP3 associating with the tetramer outer segments. It may well be that TIMP3 is the extra material represented in (E). (H) Collagen VI dimers. (I) Linear interaction of two dimers through the dimer outer segments to form a chain. (J) Lateral associations of chains of dimers with tetramers. Extra material associating with the tetramer outer segments may explain the greater electron density seen in correspondence of the main pairs of bands. (Figure from Knupp and Squire, 2003).

The different aspects that collagen VI networks present in the eyes of patients suffering from different ocular diseases suggest that collagen VI can interact with a variety of protein and retinal components (Fig. 4D). Recent immunolabeling experiments of collagen VI networks in the trabecular lamellae of the corneoscleral meshwork found that these networks contain, in addition to collagen VI, both fibrillin and decorin (Ueda and Yue, 2003). *In vitro* experiments of purified collagen VI molecules showed that, in the presence of byglycan, they can assemble into hexagonal networks (Wiberg *et al.*, 2001). The ability of collagen VI to interact with other matrix molecules is shared with collagen IV in the basement membrane, which is also able to bind to a number of basement membrane components.

C. *Type VIII Collagen*

Type VIII collagen is found in Descemet's membrane, the basement membrane that separates corneal endothelial cells from the stroma. In addition, it is also synthesized by vascular endothelial cells, and by epithelial and mesenchymal cells of other tissues (Ayad *et al.*, 1994). Two alpha chains, $\alpha 1$(VIII) and $\alpha 2$(VIII), have been seen and it has been suggested that, in Descemet's membrane, the molecule exists in the $[\alpha 1(\mathrm{VIII})]_2[\alpha 2(\mathrm{VIII})]$ form (Furthmayr *et al.*, 1983; Shuttleworth, 1997). However, *in vitro* studies showed that homotrimers made of either three $\alpha 1$(VIII) or three $\alpha 2$(VIII) chains can be formed and are stable. During the same experiments, the $[\alpha 1(\mathrm{VIII})][\alpha 2(\mathrm{VIII})]_2$ form was also created (Illidge *et al.*, 2001; Sutmuller *et al.*, 1997). The variation in the chain composition of type VIII collagen may have subtle functional differences in tissues (Shuttleworth, 1997). The triple-helical domains of the $\alpha 1$(VIII) and $\alpha 2$(VIII) chains contain eight imperfections in the Gly-X-Y repeat. Early biosynthetic studies showed that secretion of type VIII collagen was independent of the requirement for prolyl hydroxylation, in contrast to the fibrillar collagens. In the microscope, metal-shadowed collagen VIII molecules from cultured rabbit corneal endothelial cells showed rodlike structures of length 132 (± 5) nm with a large globular domain at one end (noncollagenous domain 1; NC1) and a smaller one at the other (NC2; Shuttleworth, 1997). The C-terminal NC1 domain crystallizes as a trimer, which is thought to help to form or stabilize triple helix formation in the molecule (Kvansakul *et al.*, 2003). The crystal structure also reveals three surface strips of partially solvent-exposed hydrophobic residues, which may play a significant role in the formation of the hexagonal supramolecular assemblies described below.

The organization of type VIII collagen that has been studied most closely is that in Descemet's membrane, where it forms a hexagonal lattice

with electron-dense nodes at the vertices of the hexagons, which have 160 (±15) nm sides (Shuttleworth, 1997). The hexagonal lattice structure of Descemet's membrane presumably represents a structural solution to the problem of creating a matrix that can resist compression and maintain an open porous structure (Ninomiya *et al.*, 1990).

Recombinant $[\alpha 1(\text{VIII})]_3$ and $[\alpha 2(\text{VIII})]_3$ molecules form highly ordered supramolecular assemblies. These assemblies may be formed by four triple-helical collagen VIII molecules that come together to form a tetrahedron via the hydrophobic patches on their C-termini. It has also been suggested that these tetrahedral structures may further associate to form hexagonal lattices, with the N-terminals of individual molecules interacting with either the N-terminals or with the triple-helical portion of molecules in other tetrahedrons (Stephan *et al.*, 2004).

Collagen VIII appears to be involved in differentiation and tissue remodeling. It appears to be secreted by rapidly proliferating cells and can be found in basement membranes. It was suggested that Collagen VIII acts as a bridge between different types of matrix molecules (Ricard-Blum *et al.*, 2000; Shuttleworth, 1997; Sutmuller *et al.*, 1997), and it has been found to stimulate smooth muscle cell migration and matrix metalloproteinase synthesis (Hou *et al.*, 2000). Collagen VIII is present in brain, placenta, heart, lung, and thymus of embryonic and neonatal murine tissue. It is also seen in vascular tissue, arterioles, and venules in muscle, heart, kidney, spleen, liver, lung, tendons, and cartilage matrix (Ricard-Blum *et al.*, 2000).

D. Type X Collagen

Type X collagen is a short-chain collagen that is regulated in time and location during fetal development. It is synthesized by hypertrophic chondrocytes during enchondral bone formation and, since the extracellular matrix surrounding chondrocytes mineralizes to be replaced by bone marrow and bone, it was suggested that collagen X may be associated with the mineralization process (Ayad *et al.*, 1994; Sutmuller *et al.*, 1997). Collagen X has been shown to be able to interact with chondrocytes and other cells, primarily via their $\alpha 2\beta 1$ integrins (Luckman *et al.*, 2003).

Type X collagen is made of three $\alpha 1(\text{X})$ chains. Its triple helix is about 132 nm in length, with a small N-terminal domain and a large C-terminal globular domain. As in the case of the collagen VIII NC1 domain, which has a very similar sequence (Kvansakul *et al.*, 2003), the crystal structure of the Type X C-terminal domain reveals on its surface three strips containing eight partially aromatic residues that, because of their apolar nature, are probably important for supramolecular aggregation (Bogin

et al., 2002). Also like collagen VIII, the triple-helical domains of the $\alpha 1$(X) chain contain eight imperfections in the Gly-X-Y repeat. The amino acid sequences of Collagen X and Collagen VIII α-chains are, in fact, so similar that hybrid molecules combining shortened $\alpha 1$(VIII) chains and $\alpha 1$(X) chains have been created (Illidge *et al.*, 2001).

Type X collagen seems to be deposited in the cartilage matrix without preprocessing (Ayad *et al.*, 1994). *In vitro*, Type X monomers are seen to aggregate and to form multimeric clusters arranged as a lattice. The lattice consists of a hexagonal array of nodules, presumably arising from the globular domains, interconnected by a filamentous network formed by the triple-helical parts of the collagen X molecules. The average distance between two nodules within the hexagonal lattice is about 100 ($\pm$15) nm, a distance shorter than the measured length of 130 nm of the triple-helical domain of a Type X collagen molecule (Kwan *et al.*, 1991). It was suggested that the short nodule-to-nodule distance may be in part because of the formation of superhelical structures among the adjacent helical domains. If this is the case, the mechanisms giving rise to supercoiling must be different from that proposed for the collagen VI segmented supercoil (Knupp and Squire, 2001; see below). In collagen VI, supercoiling is driven by patches of hydrophobic amino acids coiling around the collagen triple helix. This is a direct consequence of the heterotrimeric nature of collagen VI, of which there is no evidence in collagen X since only one $\alpha 1$(X) chain has been discovered so far. Lateral interactions between helical domains within individual multimeric clusters were also observed (Kwan *et al.*, 1991).

E. Dogfish Egg Case Collagen

Evidence for the collagenous nature of the dogfish egg case came from X-ray diffraction studies that indicated the existence of a 2.9 Å meridional arc typical of collagens, the thermal shrinkage characteristic curve (S-shaped with a mean half shrinkage temperature of 78 °C), and amino acid composition analysis showing that glycines accounted for about 16% of the amino acid residues (Knight and Hunt 1974; Rousaouen *et al.*, 1976). The dogfish egg case collagen was purified and observed in the electron microscope after metal shadowing. It appears to be about 45 nm long with a large globular domain at one end (about 4 nm in diameter) and a much smaller one at the other (Knight *et al.*, 1996). The short molecular length of the dogfish egg case collagen is thought to contribute to its ability to form liquid crystalline phases and crystalline networks.

Luong *et al.* (1998) sequenced, at least partially, the dogfish egg case collagen protein extracted from the nidamental gland and from newly

formed eggs. From their study, the dogfish egg collagen appears to form from two peptides in which repeated Gly-X-Y regions occur. The collagenous peptides, 34 and 36 kDa respectively, shared the same N-terminal sequence and might be two variants with different α-chain compositions. Homologies in sequence with Type IV and Type X collagen were recognized.

Before the formation of the egg case wall, the dogfish egg case collagen is stored and secreted by specialized glands (Knight *et al.,* 1996). The collagen assembles within the Golgi apparatus and migrates through the cytoplasm within storage granules. These undergo maturational changes as the collagen molecule aggregation state varies through different liquid crystalline phases. Finally, it flows through secretory ducts to the spinnerets, where it starts to form fibrils (Knight *et al.,* 1993, 1996). The liquid crystalline changes of the collagen molecules in the storage granules are thought to be driven by pH variations generated by proton pumps and by changes in the water content (Feng and Knight, 1994; Knight *et al.,* 1996). These phases have been classified by Knight *et al.* (1993) as a poorly ordered micellar phase (Phase I), a transversely banded lamellar phase (Phase II), a cholesteric mesophase without transverse banding (Phase III), a hexagonal-columnar phase (Phase IV), a second poorly oriented micellar phase (Phase V), and a highly ordered arrangement found in the mature egg case (Phase VI). *In vitro* studies carried out at pH 8, which presumably is close to the physiological pH of the egg case, generated granules showing material containing only three phases: Phase IV (hexagonal columnar; Fig. 5A), Phase VI (the mature egg arrangement; Fig. 6C), and a new arrangement with centrosymmetrically banded fibrillar material showing a period of 17.5 nm (Phase VII; Feng and Knight, 1994).

A three-dimensional reconstruction was carried out for two of these different molecular arrangements: the hexagonal columnar phase (Phase IV; Knupp *et al.,* 1999) and the arrangement in the mature egg case (Knupp *et al.* 1996, 1998). In the hexagonal columnar phase (Fig. 5A), the collagen presents good conformational order if seen in transverse view, but poorer order in longitudinal views. The hexagons seen in cross-section (Fig. 5B) measure about 20 nm on each side. The corners appear more dense than the rest (chevron) and an area of protein density occurs in the middle of the hexagons (arrowhead). Strands of protein density are seen joining the corners and the middle of each hexagon. The reconstructed images show axial columns of protein density lying regularly on the vertices of hexagonal cells (Fig. 5C, D). These columns are connected to the three nearest neighbors by sheets of protein. Distinguishable protein strands within sheets show a preferred molecular direction of about 40 to 50 degrees to the long axis (Fig. 5C, arrows). The packing scheme

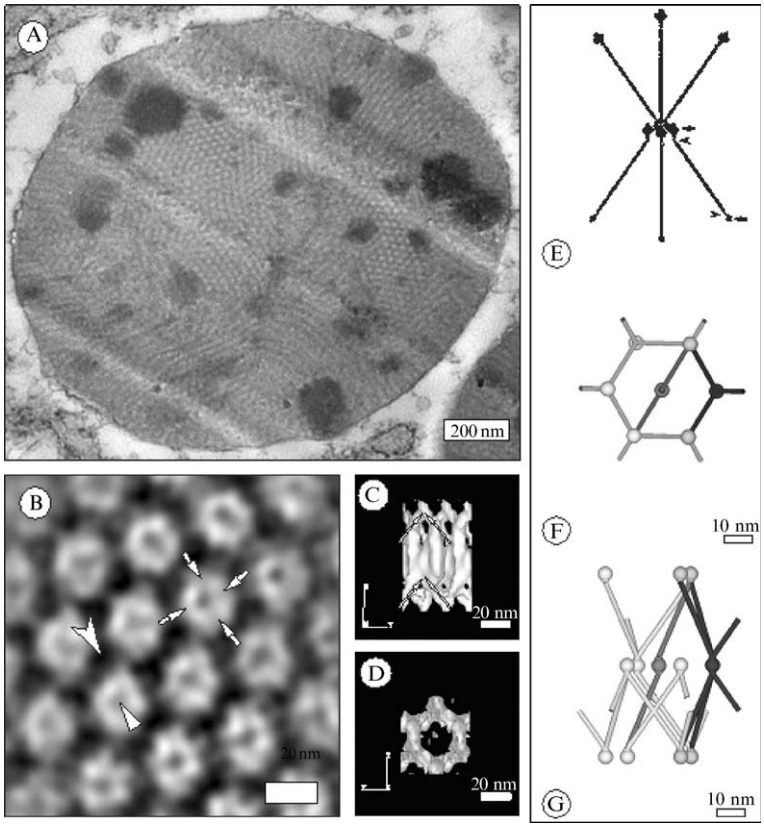


FIG. 5. Hexagonal columnar liquid crystals made of dogfish egg case collagen. (A) Typical view of a storage granule. (B) Fourier-filtered image of a transverse view. The corner of each hexagon appears relatively dense and massive (chevron), a lower protein density is seen in the center of each hexagon (arrowhead). Spikes of protein density irradiate irregularly from the center of each hexagon (arrows). (C) Longitudinal view of the three-dimensional reconstruction of the hexagonal columnar collagen arrangement. Protein strands show a preferred molecular direction of about 40 to 50 degrees to the long axis. (D) Transverse view. Axial columns of protein density lie on the vertices of hexagonal cells and are connected by irregular sheets of proteins. (E) Basic unit of the collagen arrangement in the hexagonal phase of dogfish egg case collagen. The heads of three collagen molecules join with the tails of three other molecules (arrows). Likely interruptions of the triple helices (chevrons) may endow the molecules with the necessary flexibility to form networks. (F) Transverse view and (G) longitudinal view of the molecular arrangement proposed in the hexagonal columnar phase. One basic unit is highlighted in dark gray. Occasionally, molecules can aggregate differently (light gray molecules) and give rise to the protein density seen in the middle of the hexagons (Figure from Knupp and Squire, 2003).

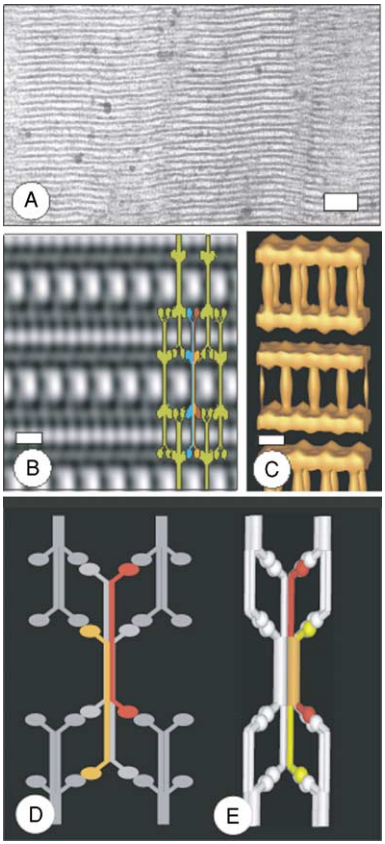


FIG. 6. (A) Typical view of the collagen arrangement within the dogfish egg case. (Bar, 100 nm.) (B) Fourier synthesis of a longitudinal view with collagen octamers superimposed. One such octamer is color-coded to highlight single monomers (red and yellow molecules) and their axial shift. A dimer is drawn in blue. (Bar, 10 nm) (C) Three-dimensional reconstruction of the collagen arrangement within the egg case. (Bar, 10 nm). (D) Two-dimensional diagram of the collagen arrangement in the dogfish egg case. Two monomers are highlighted in red and yellow, and their laterally interacting parts in orange. Monomers associate laterally to form octamers, with four molecules axially shifted by 15 nm. Interactions between different octamers occur via their globular domains. (E) As in (A), but in three dimensions. (Figure from Knupp and Squire, 2003).

proposed for the collagen molecules in the hexagonal columnar phase was based on the interaction of the three α-chains forming the molecule (Fig. 5E–G, dark gray molecules in F and G). The heads of three molecules join together with the tails of three other molecules presumably through

hydrophobic interactions (Fig. 5E, arrows). Likely interruptions of the collagen Gly-X-Y sequence allow the molecule to kink (chevrons) and assume a 40 to 50 degree tilt to the long axis. Occasionally some molecules may aggregate differently and cross the hexagonal unit cell, giving rise to the areas of protein density in the middle of the hexagons (Fig. 5F, G, gray molecules).

The arrangement of the molecules in the mature egg case (Fig. 6A) is very different from that described above. In longitudinal sections, pairs of bands of protein density about 15 nm apart are repeated axially approximately every 45 nm (Fig. 6B, C). Filaments of protein density run axially through the bands. In some views, the filaments that are separated laterally by 10 nm in the big gap between pairs of bands are staggered laterally by 5 nm when traversing the pairs of bands. In the small gap between bands, the lateral periodicity of the filaments is 5 nm. Cross-sectional views of the egg case show that the filaments are arranged on a body-centered square lattice with 10 nm sides. Consideration of the unit cell symmetry suggests that each filament in the big gap between bands is made of eight molecules, with four molecules axially shifted by ~15 nm with respect to the other four (Fig. 6D, E). The globular domains at the extremities of the octamers interact with other octamer globular domains and, in projection, give rise to the transverse bands. An axial overlap of the octamers gives rise to the doubled lateral periodicity of the filaments in the small gap between the bands in longituginal view. In cross-section, groups of octamers are arranged on a square lattice with 10 nm sides; octamers belonging to different levels are laterally shifted, so that those at one axial level are in the middle of the square lattice of the levels above and below. This arrangement is extraordinarily similar to that seen in the collagen VI ocular aggregates described above.

III. Intertwining of Network Collagens: The Type VI Segmented Supercoil

As discussed above, the fact that Type VI collagen contains more than one chain type makes it possible for the triple-helical regions of the three chains in this collagen to aggregate systematically to form a segmented supercoil (Knupp and Squire, 2001). Since it is possible that other network collagens can adopt a similar fold (e.g., along parts of Type IV collagen), the suggested origins of this supercoiling are briefly outlined here.

Figure 7 shows the basis of the analysis by Knupp and Squire (2001). Study of the properties of the central 75 nm domain was carried by separately analyzing the contributions of the positively and negatively charged residues, the hydrophobic residues, and also the numerous

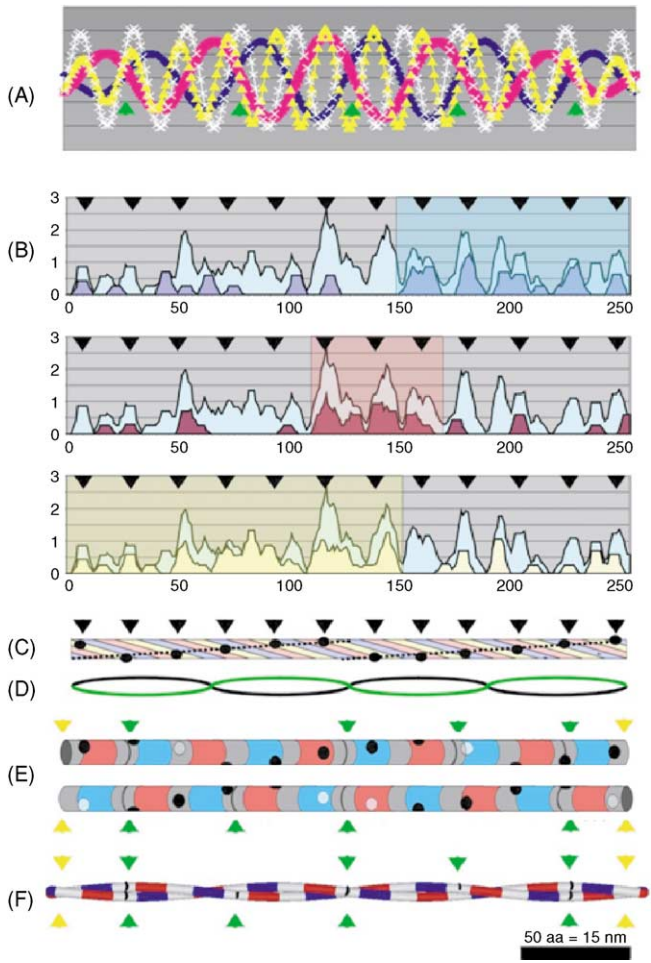


FIG. 7. Sequence character along the Type VI collagen overlap region and illustration of the segmented-coil model of dimer formation (Knupp and Squire, 2001). (A) Fourier syntheses of the polar, hydrophobic, and proline residue distributions in the 75 nm long region of the collagen VI dimer using the principal peaks in the Fourier transform of those particular amino acids in the sequence. Code: blue, the charge distribution along a molecule; red, the same charge distribution reversed as for the second (antiparallel) molecule; white, the proline distribution of each molecule; yellow, the hydrophobic distribution of each molecule. The proline and hydrophobic distributions are symmetrical so the illustrated distributions apply to both members of the antiparallel pair. The distribution of polar residues is asymmetric and complementary across the midline so that areas of high positive charge density along one molecule (blue line) are in register with minima (areas of high negative charge density) of the other molecule (red line). Charge/proline distributions and

prolines in positions X and Y of each triplet. In all cases, a property score was either determined for each chain individually or was summed over the three chains for further analysis, the latter being termed the α123 molecule. The total amino acid array length was about 250 residues. It was found that the summed polar amino acid scores for the three α-chains gave a main peak in the Fourier transform at 46.5 amino acids, with a second weaker peak at about 36.5 residues. The summed proline scores for the three α-chains had a main repeat at 23.3 amino acids, and the summed hydrophobic amino acid profile repeated at 21.3 amino acids. The discontinuities in the tripeptide sequences (in other words, the absence of glycine at the expected position, or the insertion or deletion of an amino acid giving a phase change in the Gly-X-Y sequence) occurred at approximate intervals of 44–48 amino acids apart from one missing

---

hydrophobic distributions are in phase towards the extremities of the 75 nm long region, but out of phase in the central part. Green arrows indicate the positions of the interruptions along the two portions of the triple helix. These generally occur in the middle of proline-rich regions. (B) Hydropathicity profiles in the 75 nm long region of the Type VI dimers. In the three panels, the curve underlying the light blue areas is the sum of the hydrophobic contribution of all three α-chains. In dark blue, red, and yellow are highlighted the individual contributions from the α1, α2, and α3 chains, respectively. The hydrophobic contributions are not evenly spread along the three chains; the main contributions are from α3 towards the N-terminal end (left), from α1 towards the C-terminal end and from α2 in the middle, with slightly overlapping contributions here from α1 and α3. These regions are highlighted by yellow, blue, and red background shading respectively. Black arrows highlight the 21–22 amino acid repeat of the total hydrophobic (light blue) distribution. (C) Radial projection of the collagen triple helix in the 75 nm long region of Type VI collagen that contributes to dimerization. This radial projection can be thought of as being produced by wrapping a piece of paper into a cylinder around the collagen triple helix, and then tracing the positions of the single α-chains onto it; the α1, α2, and α3 chains are shown in blue, red, and yellow, respectively. The black circles represent the locations of the main hydrophobic regions on the triple helix (their repeat is indicated by the black arrows). These amino acids lie on a track corresponding to a left-handed superhelix of pitch 37.5 nm. (D) Two antiparallel molecules (one black and the other green) can only bring together the hydrophobic patches along the corresponding two superhelices if the molecules twist around each other. (E) Diagram of the 75 nm long region contributing to dimerization on two antiparallel molecules. Red and blue regions show bands of net +ve or −ve charge; gray patches show regions of high proline density. Green arrows highlight regions where the Gly-X-Y sequence is interrupted; yellow arrows show the cystines. The positions of the main hydrophobic patches are shown by ovals around the triple helix (black in front, gray behind). (F) Model of the collagen VI segmented-supercoil showing the complementary polar interaction of the molecular bands illustrated in (E) using the same colors. The hydrophobic patches in (E) are now buried on the central line of contact of the two molecules (Bar, 50 amino acids or 15 nm, applies to all parts of the Figure).

discontinuity. Clearly, most of these repeats (polar 46.5, proline 23.3, disruptions 44–48) are closely related; only that of the hydrophobic residues (21.3) does not fit in.

Fourier syntheses of the polar, hydrophobic, and proline residue distributions in the 75-nm long region of the collagen VI dimer obtained using the principal peaks in the Fourier transform of those particular amino acids in the sequence are illustrated in Fig. 7. Taking the first two 23.3 amino acid proline repeats at the N-terminus as a starting point (Fig. 7A), the negative charges dominate in the first of these repeats (proximal to the N-terminus) and the positive charges dominate in the second. This pattern then repeats along the molecule. The sequence discontinuities occur regularly in the middle of every second proline repeat. The hydrophobic repeat is such that it starts at the N-terminal end (left hand edge of Fig. 7A) more or less in phase with the proline repeat, but then the hydrophobics get gradually out of step with the prolines, so that they are in antiphase halfway along the molecule and are back exactly in phase at the right hand (C-terminal) end. In summary, the hydrophobic pattern and the proline pattern have approximate mirror symmetry around the center of this sequence, whereas the distribution of charges is asymmetric and complementary.

Taking the amino acid distribution in one molecule and trying to match against it the appropriate features of the adjacent molecule in the pair gave rise to the following logic. First, if the interacting molecules are parallel, they must be axially shifted to produce complementary charge interactions, but then hydrophobic residues with a different repeat will inevitably be out of step in the two molecules and the cystine residues will also not line up. On the other hand, if the interacting molecules are antiparallel, then to produce complementary charge interactions they could either be overlapped by 75 nm as in Fig. 7D and E, or they could be stepped in either direction by a small multiple of the 48 polar amino-acid repeat. However, if the molecules are not overlapped by 75 nm, then the hydrophobic patches are no longer in optimal register. In fact, the only structure that brings opposite charges into alignment and also at the same time optimizes apolar interactions is the 75 nm overlapped antiparallel dimer. This also brings the cystine residues at the ends of the tripeptide regions of the two chains into axial alignment.

Further analysis of the separate chains shows that the $\alpha 1$ chain contains a higher than average number of negative charges, the $\alpha 3$ chain contains a higher than average number of positive charges, and the $\alpha 2$ chain is close to neutral. When analyzing the hydrophobic residues in individual chains (Fig. 7B), it was found that the main contribution to the total molecular hydropathicity towards the N-terminal end is from the $\alpha 3$ chain, at the

C-terminal end is from the $\alpha 1$ chain, and in the middle is from $\alpha 2$, with slightly overlapping contributions from $\alpha 1$ and $\alpha 3$. If the amino acid characteristics had been evenly spread along all three chains, then there would have been no evidence for supercoiling. Assuming that the molecular interactions are stabilized primarily by apolar interactions, the fact that the sites of greatest hydropathicity follow roughly helical tracks along the molecules (Fig. 7C–E), and that they shift from one chain to another, suggests that supercoiling must occur. The $\alpha 123$ molecule shown in Fig. 7F would give a supercoil of pitch about 37.5 nm (one half of 75 nm).

Since the sequence discontinuities occur precisely in step with the proline and polar amino acid repeats, one can imagine a pseudo-repeating unit comprising: [break–short proline region–negative patch–long proline region–positive patch–short proline region–break]. However, the fact that the hydrophobic repeat does not coincide with this main structural sequence means that, in terms of their apolar character, every one of these repeats is different. The presence of the repeats separated by discontinuities gives the supercoil a clearly segmented character, hence the preference for the name *segmented supercoil* for this collagen structure. This segmentation is also emphasized by the fact that regions of high proline content are likely to be rather rigid, whereas the polar-rich regions are likely to be more flexible. The curvature along the molecules would not be expected to be constant, but to vary slightly within each segment.

Finally, note that, as in all other supercoiled molecules (e.g., the α-helical coiled coil), the sense of the twist changes from the Type VI molecular supercoil (left-handed) to the supercoiling of the individual collagen chains in the 10/3 helix (right-handed) to the twist within each chain (left-handed). This is characteristic of ''regular lay'' rope structures and produces strength and rigidity in the assembly.

IV. Common Structural Themes in Network-forming Collagens

The different members of the network-forming collagen family vary considerably with respect to their amino acid sequence and dimensions. Nevertheless, it is possible to recognize common structural patterns in the way they assemble. Figure 8 summarizes various structural characteristics of collagen IV, VI, and dogfish egg case collagen. These three kinds of collagen are all capable of linearly associating via their globular domains (Fig. 8A–C, and axially through their triple-helical chains (Fig. 8D–F).

Collagen IV associates linearly, end-to-end, via its C-termini (C-terminus with C-terminus), while collagen VI and dogfish collagen appear to form this kind of association via a heterotypic aggregation of their globular

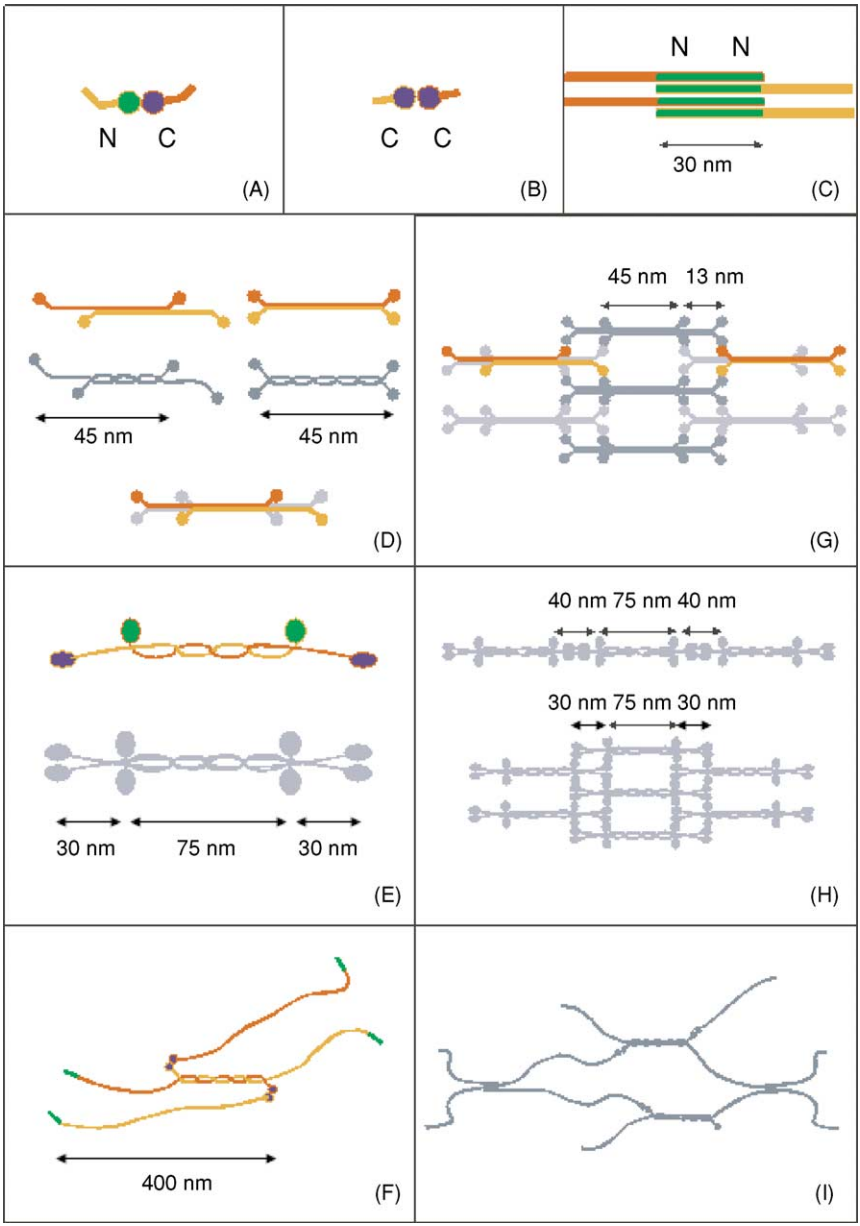


FIG. 8. Summary of the structural characteristics of collagen IV, VI, and dogfish egg case collagen. These collagens can interact linearly end-to-end through their N- and C-termini (represented in green and purple, respectively). Dogfish egg case and

domains (C-terminus with N-terminus). Further, the three types of collagen associate laterally by antiparallel interactions. Type IV collagen is also seen to associate via its 30-nm-long N-terminals (Fig. 8C) and, in addition, by intertwining of two or more collagen molecules (Fig. 8F). Collagen VI molecules associate laterally to form antiparallel dimers and tetramers (Fig. 8E). Collagen VI monomers are axially shifted by about 30 nm when forming dimers and, as discussed in Section III, they also intertwine around each other, in this case forming a segmented supercoil (Fig. 8E).

Finally, the dogfish egg case collagen associates axially into octamers. As with collagen VI, this association involves an axial displacement of four monomers by about 15 nm with respect to the other four (Fig. 8D). Although at the resolution achieved in the structural studies of the egg case carried out so far, it was not possible to distinguish any molecular supercoiling, it cannot be excluded that some kind of molecular intertwining is formed also in this case, since the resulting structure would presumably possess a functionally desirable increase in strength.

Note that a clear implication of the analysis of the segmented supercoil structure is that it will not occur if the three chains in the molecule are identical (e.g., Type X collagen). It may well be possible to generate an analogous supercoil with two chain types (e.g., the Type IV versions: $[\alpha 1(\mathrm{IV})]_2[\alpha 2(\mathrm{IV})]$ or $[\alpha 5(\mathrm{IV})]_2[\alpha 6(\mathrm{IV})]$, and Type VIII version: $[\alpha 1(\mathrm{VIII})]_2[\alpha 2(\mathrm{VIII})]$) or in parts of the three chain version of Type IV $[\alpha 3(\mathrm{IV})][\alpha 4(\mathrm{IV})][\alpha 5(\mathrm{IV})]$, but this has yet to be demonstrated.

The way collagens VIII and X assemble has not been studied in detail yet, but it appears that they too form molecular assemblies via globular interactions that involve more than two molecules. In particular, because the Type X collagen hexagonal arrangement does not account for the dimensions of the molecules (they have a smaller periodicity than expected), it was suggested that there might be axial overlapping of molecules. Dogfish egg case collagen forms a hexagonal structure as well, and

---

collagen VI interactions are represented in (A), and collagen IV interactions in (B) and (C). In addition, they all interact laterally to form antiparallel dimers. In (D), the dogfish egg collagen interactions are represented. At present, it is not known whether monomers associate with a 15 nm axial shift to form dimers, or if an axial shift occurs between already formed tetramers. Although the resolution reached does not allow a direct visualization, it is conceivable that supercoiling occurs between monomers. Lateral interactions of collagen VI are represented in (E) and those of collagen IV in (F). Networks are formed when linear and lateral interactions take place: the dogfish egg case collagen network is represented in (G), the collagen VI network in (H), and the collagen IV network in (I). Collagen VI is also capable of forming long chains when tetramers associate linearly (H, top).

it is likely that hexagonal symmetry is a consequence of the intrinsic threefold character of the collagen molecules due to their triple-helical conformation.

Linear and axial interactions of the kind described above appear fundamental in the formation of collagen networks. The interactions formed by C- and N-terminals help to form linear successions of molecules, while the axial interactions are important in the generation of open networks since they involve molecules belonging to different levels. In addition, intertwining of different collagen molecules may be another common theme. This behavior is not present in other collagens (e.g., collagen I). Interruptions in the Gly-X-Y sequence of the network-forming collagen is a common characteristic. This may help to reduce the stiffness of the collagen molecule, allowing more spatial freedom for the molecule, and may also promote supercoiling. Finally, collagen IV and VI seem to be able to interact with many other proteins and macromolecules. It is probable that this characteristic is used in strengthening the networks formed by these collagens, as well as having other physiological implications.

Acknowledgments

The authors wish to thank the Wellcome Trust for their financial support (Grant # 389048) and Dr. Michael Chew for Fig. 1A.

References

Ayad, S., Boot-Hanford, R. P., Humphries, M. J., Kadler, K. E., and Shuttleworth, C. A. (1994). ''The Extracellular Matrix Facts Book.'' Academic Press, London, San Diego.

Baldock, C., Sherrat, M. J., Shuttleworth, C. A., and Kielty, C. M. (2003). The supermolecular organisation of collagen VI microfibrils. *J. Mol. Biol.* **330,** 297–307.

Bogin, O., Kvansakul, M., Rom, E., Singer, J., Yayon, A., and Hohenester, E. (2002). Insight into Schmid metaphyseal chondrodysplasia from the crystal structure of the collagen X NC1 domain trimer. *Structure* **10,** 165–173.

Ball, S., Bella, J., Kielty, K., and Shuttleworth, A. (2003). Structural basis of type VI collagen dimer formation. *J. Biol. Chem.* **278,** 15326–15332.

Borza, D. B., Bondar, O., Ninomiya, Y., Sado, Y., Naito, I., Todd, P., and Hudson, B. G. (2001). The Nc1 domain of collagen IV encodes a novel network composed of the $\alpha1$, $\alpha2$, $\alpha5$ and $\alpha6$ chains in smooth muscle basement membranes. *J. Biol. Chem.* **276,** 28532–28540.

Boutaud, A., Borza, D. B., Bondar, O., Gunwar, S., Netzer, K. O., Singh, N., Ninomiya, Y., Sado, Y., Noelken, M. E., and Hudson, B. G. (2000). Type IV collagen of the glomerular basement membrane. *J. Biol. Chem.* **275,** 30716–30724.

Brazel, D., Pollner, R., Oberbaumer, I., and Kuhn, K. (1988). Human basement membrane collagen Type IV: The amino acid sequence of the $\alpha2$(IV) chain and

its comparison with the $\alpha 1$(IV) chain reveals deletions in the $\alpha 1$(IV) chain. *Eur. J. Biochem.* **172,** 35–42.

Butkowsky, R. J., Langeveld, J. P., Wieslander, J., Hamilton, J., and Hudson, B. G. (1987). Localization of the Goodpasture epitope to a novel chain of basement membrane collagen. *J. Biol. Chem.* **262,** 7874–7877.

Camacho Vanegas, O., Bertini, E., Zhang, R. Z., Petrini, S., Minosse, C., Sabatelli, P., Giusti, B., Chu, M. L., and Pepe, G. (2001). Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen Type VI. *Proc. Natl. Acad. Sci. USA* **98,** 7516–7521.

Engel, U., Pertz, O., Fauser, C., Engel, J., David, C. N., and Holstein, T. W. (2001). A switch in disulfide linkage during mini-collagen assembly in hydra nematocyst. *EMBO J.* **20,** 3063–3073.

Feng, D., and Knight, D. P. (1994). The effect of pH on fibrillogenesis of collagen in the egg capsule of the dogfish, *Scyliorhinus Canicula. Tissue Cell* **26,** 649–659.

Furthmayr, H., Wiedemann, H., Timpl, R., Odermatt, E., and Engel, J (1983). Electron microscopical approach to a structural model of intima collagen. *Biochem. J.* **211,** 303–311.

Gunwar, S., Ballester, F., Noelken, M. E., Sado, Y., Ninomiya, Y., and Hudson, B. G. (2001). Glomerular basement membrane: Identification of a novel disulfide-cross-linked network of $\alpha 3$, $\alpha 4$, and $\alpha 5$ chains of Type IV collagen and its implication for the pathogenesis of Alport syndrome. *J. Biol. Chem.* **273,** 8767–8775.

Holstein, T. W., Benoit, M., Herder, G. V., Wanner, G., David, C. N., and Gaub, H. E. (1994). Fibrous mini-collagens in hydra nematocysts. *Science* **256,** 402–404.

Hou, G., Mulholland, D., Gronska, M. A., and Bendeck, M. P. (2000). Type VIII collagen stimulates smooth muscle cell migration and matrix metalloproteinase synthesis after arterial injury. *Am. J. Pathol.* **156,** 467–476.

Illidge, C., Kielty, C., and Shuttleworth, A. (2001). Type VIII collagen: Heterotrimeric chain association. *Int. J. Biochem. C. Biol.* **33,** 521–529.

Kahsai, T. Z., Enders, G. C., Gunwar, S., Brunmark, C., Wieslander, J., Kalluri, R., Zhou, J., Noelken, M. E., and Hudson, B. G. (1997). Seminiferous tubule basement membrane. Composition and organisation of type IV collagen chains, and the linkage of $\alpha 3$(IV) and $\alpha 5$(iv) chains. *J. Biol. Chem.* **272,** 17023–17032.

Kalluri, R., and Cosgroves, D. (2000). Assembly of type IV collagen: Insights from $\alpha 3$(IV) collagen-deficient mice. *J. Biol. Chem.* **275,** 12719–12724.

Kielty, C. M., Whittaker, S. P., Grant, M. E., and Shuttleworth, C. A. (1992). Type VI collagen microfibrils: Evidence for a structural association with Hyaluronan. *J. Cell Biol.* **118,** 979–990.

Knight, D. P., and Hunt, S. (1974). Fibril structure of collagen in egg capsule of Dogfish. *Nature* **49,** 379–380.

Knight, D. P., Feng, D., Stewart, M., and King, E. (1993). Changes in macromolecular organisation in collagen assemblies during secretion in the Nidamental Gland and formation of the egg capsule wall in the Dogfish *Scyliorhinus Canicula. Phil. Trans. R. Soc. Lond. B* **341,** 419–436.

Knight, D. P., Feng, D., and Stewart, M. (1996). Structure and function of the selachian egg case. *Biol. Rev.* **71,** 81–111.

Knupp, C., and Squire, J. M. (2001). A new twist in the collagen story: The Type VI segmented supercoil. *EMBO J.* **20,** 372–376.

Knupp, C., and Squire, J. M. (2003). Molecular packing in network-forming collagens. *Sci. World J.* **3,** 558–577.

Knupp, C., Chew, M. W. K., Morris, E. P., and Squire, J. M. (1996). Three-dimensional reconstruction of a collagen IV analogue in the dogfish egg case wall. *J. Struct. Biol.* **117,** 209–221.

Knupp, C., Chew, M. W. K., and Squire, J. M. (1998). Collagen packing in the dogfish egg case wall. *J. Struct. Biol.* **122,** 101–110.

Knupp, C., Luther, P. K., Morris, E. P., Knight, D. P., and Squire, J. M. (1999). Partially systematic molecular packing in the hexagonal columnar phase of dogfish egg case collagen. *J. Struct. Biol.* **126,** 121–130.

Knupp, C., Munro, P. M. G., Luther, P. K., Ezra, E., and Squire, J. M. (2000). Structure of abnormal molecular assemblies (Collagen VI) associated with Human Full Thickness Macular Holes. *J. Struct. Biol.* **129,** 38–47.

Knupp, C., Amin, S. Z., Munro, P. M. G., Luthert, P. J., and Squire, J. M. (2002a). Collagen VI assemblies in age-related macular degeneration. *J. Struct. Biol.* **139,** 181–189.

Knupp, C., Chong, N. H. V., Munro, P. M. G., Luthert, P. J., and Squire, J. M. (2002b). Analysis of the collagen VI assemblies associated with Sorsby's Fundus Dystrophy. *J. Struct. Biol.* **137,** 31–40.

Kuo, H-J., Keene, D. R., and Glanville, R. W. (1995). The macromolecular structure of Type VI collagen formation and stability of filaments. *Eur. J. Biochem.* **232,** 364–372.

Kvansakul, M., Bogin, O., Hohenester, E., and Yayon, A. (2003). Crystal structure of the collagen alpha1(VIII) NC1 trimer. *Matrix Biol.* **22,** 145–152.

Kwan, A. P. L., Cummings, C. E., Chapman, J. A., and Grant, M. E. (1991). Macromolecular organisation of chicken Type X collagen in vitro. *J. Cell Biol.* **114,** 597–604.

Luckman, S. P., Rees, E., and Kwan, A. P. (2003). Partial characterization of cell-Type X collagen interactions. *Biochem. J.* **372,** 485–493.

Luong, T-T., Boutillon, M-M., Garrone, R., and Knight, D. P. (1998). Characterisation of selachian egg case collagen. *Biochem. Biophys. Res. Com.* **250,** 657–663.

Luse, S. A. (1960). Electron microscopic studies of brain tumors. *Neurology* **10,** 881–905.

Ninomiya, Y., Castagnola, P., Gerecke, D., Gordon, M. K., and Jacenko, O. (1990). The molecular biology of collagens with short triple helical domains. *In* ''Extracellular Matrix Genes,'' (Sandell and Boyd, Eds.). Academic Press, New York.

Oohashi, T., Sugimoto, M., Mattei, M.-G., and Ninomiya, Y. (1994). Identification of a new collagen IV chain, α6(IV), by cDNA isolation and assignment of the gene to chromosome Xq22, which is the same locus for COL4A5. *J. Biol. Chem.* **269,** 7520–7526.

Ozbek, S., Engel, U., and Engel, J. (2002). A switch in disulfide linkage during mini-collagen assembly in hydra nematocysts or how to assemble a 150-bar-resistant structure. *J. Struct. Biol.* **137,** 11–14.

Pokidysheva, E., Milbradt, A. G., Meier, S., Renner, C., Haussinger, D., Bachinger, H. P., Moroder, L., Grzesiek, S., Holstein, T. W., Ozbek, S., and Engel, J. (2004). The structure of the Cys-rich terminal domain of Hydra mini-collagen, which is involved in disulfide networks of the nematocyst wall. *J. Biol. Chem.* **279,** 30395–30401.

Reale, E., Groos, S., Luciano, L., Eckardt, C., and Eckardt, U. (2001). In the Mammalian Eye Type VI collagen forms three morphologically different aggregates. *Matrix Biol.* **20,** 37–51.

Ricard-Blum, S., Dublet, B., and van der Rest, M. (2000). ''Unconventional Collagens. Types VI, VII, VIII, IX, X, XII, XIV, XVI, and XIX.'' Oxford University Press, Oxford.

Rousaouen, M., Pujol, J. P., Bocquet, J., Veillard, A., and Borel, J. P. (1976). Evidence of collagen in the egg capsule of the dogfish, *Scyliorhinus canicula*. *Comp. Biochem. Physiol.* **53B,** 539–543.

Sado, Y., Kagawa, M., Naito, I., Ueki, Y., Seki, T., Momota, R., Oohashi, T., and Ninomiya, Y. (1998). Organization and expression of basement membrane collagen IV genes and their roles in human disorders. *J. Biochem. (Tokyo)* **123,** 767–776.

Scacheri, P. C., Gillanders, E. M., Subramony, S. H., Vedanarayanan, V., Crowe, C. A., Thakore, N., Bingler, M., and Hoffman, E. P. (2002). Novel mutations in collagen VI genes: Expansion of the Bethlem myopathy phenotype. *Neurology* **58,** 593–602.

Shuttleworth, C. A. (1997). Type VIII collagen. *Int. J. Biochem. Cell Biol.* **29,** 1145–1148.

Siebold, B., Qian, R., Glanville, R. W., Hofmann, H., Deutzmann, R., and Khun, K. (1987). Construction of a model for the aggregation and cross linking region (7S domain) of Type IV collagen based upon an evaluation of the primary structure of the alpha1 and alpha2 chains in this region. *Eur. J. Biochem.* **168,** 569–575.

Siebold, B., Deutzmann, R., and Khun, K. (1988). The arrangement of inter- and intramolecular disulphide bonds in the carboxyterminal, non-collagenous aggregations and cross-linking domain of basement membrane Type IV collagen. *Eur. J. Biochem.* **176,** 617–624.

Skaer, R. J., and Picken, L. E. R. (1965). The structure of the nematocyst thread and the geometry of discharge in *corynactis viridis* allman. *Phil. Trans. Roy. Soc.* **250,** 131–164.

Stephan, S., Sherratt, M. J., Hodson, N., Shuttleworth, C. A., and Kielty, C. M. (2004). Expression and supramolecular assembly of recombinant alpha1(VIII) and alpha2(VIII) collagen homotrimers. *J. Biol. Chem.* **279,** 21469–21477.

Sundaramoorthy, M., Meiyappan, M., Todd, P., and Hudson, B. G. (2002). Crystal structure of NC1 domains: Structural basis for type IV collagen assembly in basement membranes. *J. Biol. Chem.* **277,** 31142–31153.

Sutmuller, M., Bruijin, J. A., and de Heer, E. (1997). Collagens Type VIII and X; two non-fibrillar, short chain collagens. Structure homologies, functions and involvement in pathology. *Histol. Histopathol.* **12,** 557–566.

Than, M. E., Henrich, S., Huber, R., Ries, A., Mann, K., Kuhn, K., Timpl, R., Bourenkov, G. P., Bartunik, H. D., and Bode, W. (2002). The 1.9-Å crystal structure of the non-collagenous (NC1) domain of human placenta collagen IV shows stabilization via a novel type of covalent Met-Lys cross-link. *Proc. Natl. Acad. Sci. USA* **99,** 6607–6612.

Timpl, R. (1989). Structure and biological activity of basement membrane proteins. *Eur. J. Biochem.* **180,** 487–502.

Timpl, R., and Brown, J. C. (1996). Supermolecular assembly of basement membranes. *Bioessays* **18,** 123–132.

Ueda, J., and Yue, B. Y. (2003). Distribution of myocilin and extracellular matrix components in the corneoscleral meshwork of human eyes. *Invest. Ophthalmol. Vis. Sci.* **44,** 4772–4779.

Vanacore, R. M., Shanmugasundararaj, S., Friedman, D. B., Bondar, O., Hudson, B. G., and Sundaramoorthy, M. (2004). The $\alpha1.\alpha2$-network of collagen IV: Reinforced stabilization of the non-collagenous domain-1 by non-covalent forces and the absence of Met-Lys crosslinks. *J. Biol. Chem.* **279,** 44723–44730.

Von Der Mark, H., Aumailley, M., Wick, G., Fleischmayer, R., and Timpl, R. (1984). Immunochemistry, genuine size and tissue localisation of Collagen VI. *Eur. J. Biochem.* **142,** 243–502.

Wiberg, C., Heinegard, D., Wenglen, C., Timpl, R., and Morgelin, M. (2001). Biglycan organizes collagen VI into hexagonal-like networks resembling tissue structures. *J. Biol. Chem.* **277,** 49120–49126.

Wu, J-J., Eyre, D. R., and Slayter, H. S. (1987). Type VI collagen of the invertebral disc. *Biochem. J.* **248,** 373–381.

Yurchenco, P. D., and Furthmayr, H. (1984). Self-assembly of basement membrane collagen. *Biochemistry* **23,** 1839–1850.

Yurchenco, P. D., and Ruben, G. C. (1987). Basement membrane structure *in situ*: Evidence for lateral associations in the type IV collagen network. *J. Cell Biol.* **105,** 2559–2568.

Yurchenco, P. D., and Schittny, J. C. (1990). Molecular architecture of basement membranes. *FASEB J.* **14,** 1577–1590.

Yurchenco, P. D., and O'Rear, J. (1993). Supermolecular Organisation of Basement Membranes. *In* ''Molecular and Cellular Aspects of Basement Membranes,'' (D. H. Rohrbach and R. Timpl, Eds.). Academic Press, New York.

FIBRILLIN MICROFIBRILS

By CAY M. KIELTY,* MICHAEL. J. SHERRATT,† ANDREW MARSON,* AND CLAIR BALDOCK*

*Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, United Kingdom;
†Division of Laboratory and Regenerative Medicine, Faculty of Medical and Human Sciences, University of Manchester, Manchester M13 9PT, United Kingdom

I. Introduction 406
II. Fibrillin Molecules 406
A. cbEGF-Like Domains 408
B. TB Motifs 409
C. Hybrid Motifs 410
D. Proline/Glycine-Rich Regions 410
E. N- and C-Termini 410
F. Molecular Organization 411
III. Fibrillin Assembly 412
A. Role of Cells 412
B. Fibrillin-1 Homotypic Interactions 412
C. Fibrillin-1 Interactions with Microfibril-Associated Molecules 414
D. Microfibril Maturation 417
IV. Microfibril Organization and Elasticity 417
A. Organization 417
B. Elasticity 419
V. Fibrillin-1 Alignment Models 420
A. Hinged Model 420
B. Staggered Model 423
VI. Fibrillin Microfibrils in Aging and Disease 426
A. Aging 426
B. Fibrillin Heritable Diseases 426
VII. Summary 429
References 429

Abstract

Fibrillin microfibrils are widely distributed extracellular matrix assemblies that endow elastic and nonelastic connective tissues with long-range elasticity. They direct tropoelastin deposition during elastic fibrillogenesis and form an outer mantle for mature elastic fibers. Microfibril arrays are also abundant in dynamic tissues that do not express elastin, such as the ciliary zonules of the eye. Mutations in fibrillin-1—the principal structural component of microfibrils—cause Marfan syndrome, a heritable disease with severe aortic, ocular, and skeletal defects. Isolated fibrillin-rich microfibrils have a complex 56 nm ''beads-on-a-string'' appearance; the molecular basis of their assembly and

DOI: 10.1016/S0065-3233(04)70012-1
Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

elastic properties, and their role in higher-order elastic fiber formation, remain incompletely understood.

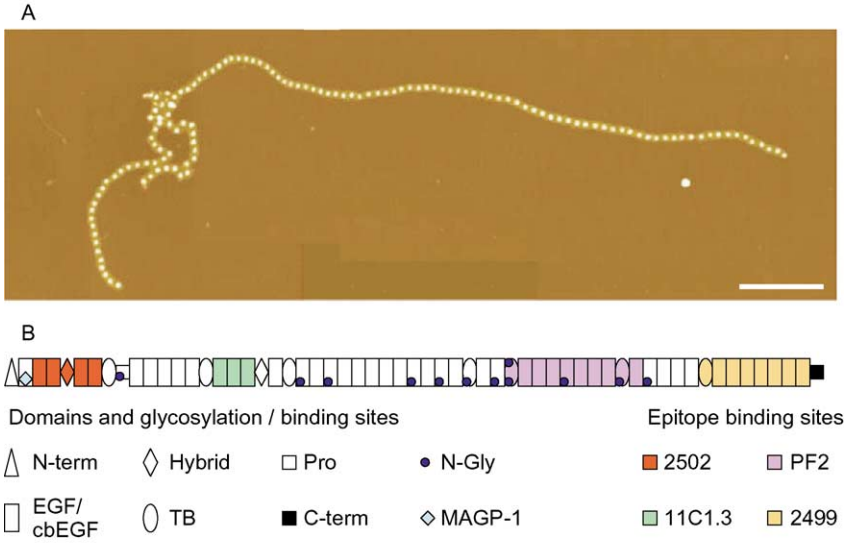
I. Introduction

Fibrillin microfibrils are widely distributed elastic microfibrils that endow connective tissues with long-range elasticity (Handford *et al.*, 2000; Kielty *et al.*, 2002a). They are evolutionarily highly conserved from jellyfish to man, which confirms their critical biomechanical importance. Microfibrils are particularly abundant in elastic tissues such as aorta, lung, and skin, where they direct tropoelastin deposition during elastic fibrillogenesis and form an outer mantle for mature elastic fibers (Kielty *et al.*, 2002b; Mecham and Davis, 1994). Microfibril arrays are also abundant in dynamic tissues that do not express elastin, such as the ciliary zonules of the eye that hold the lens in dynamic suspension (Ashworth *et al.*, 2000). Structural analysis of isolated fibrillin-rich microfibrils has revealed a complex 56 nm ''beads-on-a-string'' appearance (Fig. 1A), while the length of a fibrillin monomer is ~160 nm (Lin *et al.*, 2002; Sherratt *et al.*, 2001). Mutations in fibrillin-1 cause Marfan syndrome (MFS), a heritable disease associated with severe aortic, ocular, and skeletal defects due to defective elastic fibers, as well as several related fibrillinopathies (Robinson and Booms, 2001).

In man, although there are three fibrillin genes, fibrillin-1 is the major structural component of microfibrils. Fibrillin molecules are large (~340 kDa) and complex multidomain glycoproteins. Fibrillin-1 and fibrillin-2 have distinct but overlapping patterns of expression (Zhang *et al.*, 1995). Fibrillin-2 is generally expressed earlier in development than fibrillin-1 and may be particularly important in elastic fiber formation (Zhang *et al.*, 1994). Fibrillin-3 was isolated from brain, and it is not known whether it forms microfibrils (Corson *et al.*, 2004; Nagase *et al.*, 2001). In this Chapter, we delineate the structural organization of fibrillin-1 molecules and microfibrils, and evaluate current models of fibrillin-1 alignment within microfibrils.

II. Fibrillin Molecules

The three closely related fibrillin isoforms have distinct, but overlapping, developmental and adult tissue distributions. Human fibrillin-1 consists of 2871 amino acids, with a calculated pI of 4.81 and a predicted molecular mass of 347 kDa (Pereira *et al.*, 1993) (Fig. 1B). There are 14 N-linked glycosylation consensus sequences that, in recombinant



C

Antibody	Epitope position Amino acids	Epitope position Exons
2502	116-329	4-8
11C1.3	723-909	18-20
PF2	1759-2145	42-52
2499	2402-2730	57-64

FIG. 1. (A) Ultrastructural appearance of isolated fetal bovine aorta fibrillin microfibrils. Using intermittent contact mode atomic force microscopy beads are identified as repeating height maxima (white) with an axial periodicity of 56 nm. (Bar, 500 nm.) (B) Domain structure of N- and C-termini processed human fibrillin-1. Consensus N-linked glycosylation sequences and the potential N-terminal MAGP-1 binding site are depicted by blue circles and a cyan diamond respectively. Domains are colored according to experimentally determined antibody binding positions. (C) Localization of fibrillin antibody epitopes by Western blotting of recombinant protein fragments.

fragments, are occupied (Bax *et al.*, 2003; Rock *et al.*, 2004). The signal peptide is located at the extreme N-terminus and consists of 27 amino acids. Fibrillin-1 contains 47 epidermal growth factor-(EGF-) like domains: 43 are calcium-binding (cbEGF)-like domains, seven are 8-cysteine (TB)

modules, two are hybrid motifs with similarities to both cbEGF-like domains and TB motifs, and there is a proline-rich region that may act as a hinge region. In fibrillin-2, this sequence is glycine-rich (Zhang *et al.*, 1994); in fibrillin-3, it is proline- and glycine-rich (Corson *et al.*, 2004). N- and C-terminal furin cleavage sites imply key roles for pericellular processing in fibrillin-1 assembly. There is extreme evolutionary conservation of fibrillin-1 among mammalian species (human, porcine, bovine, murine). There is also remarkable nucleotide identity within the 5′ flanking sequence, suggesting a regulatory role for this region.

A. *cbEGF-Like Domains*

EGF-like domains are a commonly identified protein module in multidomain proteins (Downing *et al.*, 1996; Smallridge *et al.*, 2003). They contain six cysteine residues that form intradomain disulphide bonds in a 1–3, 2–4, and 5–6 manner. A subset of these domains contains a calcium binding consensus sequence of $(D/N)X(D/N)(E/Q)X_m(D/N)^*X_n(Y/F)$, where m and n are variable, and the asterisk indicates a potential β-hydroxylation site (here designated cbEGF-like domains). cbEGF-like domains often exist as tandem repeats, and calcium binding is critical for their structural integrity and function (Downing *et al.*, 1996; Werner *et al.*, 2000). More than 60% of the mutations causing Marfan syndrome identified to date occur within fibrillin-1 cbEGF-like domains, emphasizing that correct folding of these domains is critical for molecular function. Calcium plays a general role in maintaining structural rigidity of the interdomain linkage, and stabilizes cbEGF-like domain pair structures (Downing *et al.*, 1996). Tandem repeats of fibrillin-1 cbEGF-like domains are predicted to form rod-shaped structures in the presence of calcium.

High-resolution structures for several fibrillin-1 cbEGF-like domains have been determined by nuclear magnetic resonance (NMR) and X-ray crystallography (Downing *et al.*, 1996; Lee *et al.*, 2004; Smallridge *et al.*, 2003). In solution, covalently linked cbEGF-like domains (cbEGF32-33) were orientated in a near-linear extended arrangement, which was stabilized by calcium ligation of the C-terminal domain and interdomain hydrophobic packing interactions. The NMR solution structure of another calcium-loaded fibrillin-1 cbEGF-like domain pair was also solved. cbEGF12-13 also exhibits a near-linear, rodlike arrangement of domains, which suggests that all fibrillin-1 cbEGF-cbEGF pairs characterized by a single interdomain linker residue possess this rodlike structure. The domain arrangement of cbEGF12-13 was stabilized by additional interdomain packing interactions to those observed for cbEGF32-33. The cbEGF12-13

domain pair is within the longest run of cbEGFs, and many mutations that cluster in this region are associated with severe, neonatal MFS (Robinson and Booms, 2001). There is a correlation between backbone flexibility and calcium binding affinity. For both tandem cbEGF domain pairs, the N-terminal domain binds calcium more weakly than the C-terminal domain. However, the affinities of the N- and C-terminal domains of cbEGF12-13 for calcium are significantly higher than those observed for cbEGF32-33, so there is also variation between domain pairs in terms of calcium affinities in relation to flexibility. Mutations in cbEGF-like domains that affect cysteine residues are likely to alter disulfide bond formation, thereby disrupting the correct fold, while mutations affecting residues in the calcium binding consensus sequence reduce calcium binding affinity, leading to structural destabilization. Other mutations could impair domain folding. Missense mutations that change calcium binding ligands cause increased proteolytic susceptibility of recombinant fibrillin fragments (Booms *et al.*, 2000; Reinhardt *et al.*, 2000a; Whiteman and Handford, 2003).

B. TB Motifs

Fibrillin TB motifs (also known as eight cysteine motifs) are so-called because of their homology to motifs found in the latent TGFβ binding proteins (LTBPs), some of which can covalently bind the cytokine TGFβ. TB motifs are found only in the fibrillin-LTBP superfamily of proteins and contain eight cysteine residues, including a contiguous internal cluster of three cysteine residues within an α-helical region. The eight cysteine residues disulphide bond in a 1–3, 2–6, 4–7, 5–8 pattern. In fibrillin-1, six of these motifs are interspersed among the cbEGF-like domains, and a seventh precedes the proline-rich region. The TB motif is a potential source of flexibility within fibrillin-1 through its interactions with flanking cbEGF-like domains. The structure of the fibrillin-1TB6 motif itself was established by NMR (Yuan *et al.*, 1997). It is globular and comprises two α-helices and six β-strands. Strands A, D, E, and F form a four-stranded β-sheet, while the B and C strands form a two-stranded β-hairpin.

The NMR study of fibrillin-1 TB6-cbEGF32 also identified a flexible linker between these domains (Yuan *et al.*, 1997). The crystal structure of a cbEGF22-TB4-cbEGF23 fragment from fibrillin-1, which contains an Arg-Gly-Asp (RGD) integrin recognition site, has also been determined in Apo (one-calcium bound) and Holo (two-calcium bound) forms (Lee *et al.*, 2004). This study revealed novel pairwise interactions mediated by TB4, one of which may contribute to fibrillin-1 elasticity (see Section V below). Fibrillin-1 TB4 is recognized by integrin receptors $\alpha5\beta1$ and $\alpha v\beta3$ in an

RGD-dependent manner (Bax *et al.*, 2003). The structure of the third TB module in LTBP-3 ($TB3_{LTBP1}$) has also been solved (Lack *et al.*, 2003). Unlike fibrillin TB motifs, $TB3_{LTBP1}$ contains a Phe-Pro insertion that renders the 2,6 disulphide bond more easily accessible for protein–protein interaction with the TGFβ1 propeptide. The 2,6 disulphide bond can form a complex with the TGFβ1 propeptide (LAPβ1) and it is not required for domain folding. Other unique features revealed from the structure of $TB3_{LTBP1}$ are two hydrophobic patches that may be involved in intra- or intermolecular protein–protein interactions, and five negatively charged residues surrounding the 2,6 disulphide bond that result in a large electrostatic surface potential.

C. Hybrid Motifs

Hybrid motifs are apparently unique to fibrillins and LTBPs. They contain primary structural features of both TB motifs towards their N-termini and cbEGF-like domains towards their C-termini. No high-resolution structures have yet been determined for these motifs. Most hybrid motifs contain eight cysteines, but the first hybrid motif of fibrillin-1 contains nine cysteines; one of these cysteines is predicted to be important in forming intermolecular disulphide bonds that stabilize assembled fibrillin-1 (see Section III below; Reinhardt *et al.*, 2000b).

D. Proline/Glycine-Rich Regions

Fibrillin-1 contains a proline-rich region (42% proline content), in contrast to fibrillin-2 (which contains a glycine-rich region) and fibrillin-3 (which has both a proline- and glycine-rich region). While there is limited sequence homology between these regions, all have the potential to fold, which may have been an evolutionary requirement for this region. No Marfan-causing mutations have been identified within the proline-rich region of fibrillin-1, which implies a critical role.

E. N- and C-Termini

The N-terminal region of fibrillin-1 (the first 29 residues after signal peptide) contains largely basic residues (estimated pI, 11.1). There is an N-terminal putative furin cleavage site at the N-terminus (RAKR↓R; residues 41–45). Recombinant secreted N-terminal fibrillin-1 can be furin processed (Raghunath *et al.*, 1999; Ritty *et al.*, 1999; Wallis *et al.*, 2003). In our mammalian recombinant expression system, a significant level of

secreted N-terminal fragments is not cleaved at this site, although this could reflect high expression. Our mass spectrometry studies of isolated microfibrils have failed to reveal any tryptic peptides from fibrillin-1 sequences upstream of the N-terminal furin cleavage site (Cain *et al.*, 2004).

The C-terminal sequence after the last cbEGF-like domain comprises 184 residues and contains a major furin cleavage site (RKRR↓; residues 2728–2732). The unique sequence prior to the cleavage site contains a CxxC sequence that could have protein disulphide isomerase (PDI) activity, and thus the ability to (re)arrange disulphide bonds. The processed C-terminal unique sequence contains three N-linked glycosylation sites just after the furin processing site, which can regulate furin processing (Ashworth *et al.*, 1999a). On pericellular furin cleavage, a 20 kDa N-glycosylated fragment is released. Furin processing appears to be essential for microfibril deposition (Raghunath *et al.*, 1999). However, using mass spectrometry analysis, we have shown that tissue-isolated microfibrils from zonules, skin, and aorta all contain at least some C-terminal fragments, as judged by detection of tryptic peptides from the post-furin cleavage site sequence (Cain *et al.*, 2005). Thus, the significance of C-terminal processing in assembly will need to be reevaluated.

F. Molecular Organization

Low-resolution imaging of fibrillin-1 molecules and molecular fragments has been achieved by rotary shadowing (Lin *et al.*, 2002; Reinhardt *et al.*, 1996a). By this approach, fibrillin-1 appears as relatively flexible rodlike molecules of ∼160 nm with evidence for several small globular domains. In the presence of calcium, fibrillin-1 molecules appear shorter with a wider diameter, and more flexible (Reinhardt *et al.*, 1997). We have used atomic force microscopy to image unshadowed fibrillin microfibrils (Fig. 1A; Kielty *et al.*, 2002a; Sherratt *et al.*, 2004), as well as recombinant fibrillin-1 fragments (unpublished data). We are currently completing the solution structure of fibrillin-1 using small angle X-ray scattering.

The amino acid sequence of fibrillin-1 is predominantly hydrophilic, even in the presence of bound calcium (Sherratt *et al.*, 2004). In addition, fibrillin-1 has 14 potential N-glycosylation sites that may contribute to the hydrophilic nature of fibrillin-1, since complex N-linked oligosaccharides are negatively charged if they contain sialic acid. Additional regions of increased hydrophilicity may be contributed by other microfibril components (see Section III below).

III. Fibrillin Assembly

Fibrillin-1 microfibrils are highly complex extracellular polymers, and the molecular basis of assembly is not well understood. Studies have focused on defining the role of cells in regulating assembly, the molecular interactions that drive assembly, and the role of microfibril-associated molecules in the assembly process.

A. *Role of Cells*

Microfibril assembly is, at least in part, a cell-regulated process that proceeds independently of tropoelastin. Fibrillin-1 may possibly undergo limited initial assembly in the secretory pathway (Ashworth *et al.*, 1999b; Trask *et al.*, 1999), and in this respect is similar to other major extracellular matrix (ECM) macromolecules such as collagens, laminins, and proteoglycans. Intracellular chaperone associations are likely to play key roles in molecular folding and N-glycosylation (Ashworth *et al.*, 1999b; Wallis *et al.*, 2003), and in preventing inappropriate intracellular assembly. Cells also regulate the pericellular location of N- and C-terminal cleavage by furin (Raghunath *et al.*, 1999; Ritty *et al.*, 1999).

Microscopy studies have shown that assembly occurs in association with cell surfaces and predict a key role for receptors in this process. A central role for receptors has also been shown for fibronectin assembly, in which dimer interactions with $\alpha5\beta1$ integrins induce a conformation change required for linear assembly (Sechler *et al.*, 2001). Since the RGD sequence in the fourth fibrillin-1 TB motif interacts with $\alpha v\beta3$ (Bax *et al.*, 2003; Lee *et al.*, 2004; Pfaff *et al.*, 1996; Sakamoto *et al.*, 1996) and $\alpha5\beta1$ (Bax *et al.*, 2003), these receptors might influence microfibril assembly in a similar manner. It has also been shown that heparan sulphate proteoglycans (HSPGs), possibly in the form of cell surface syndecan receptors, may have a role in assembly (Tiedemann *et al.*, 2001). Pericellular chondroitin sulphate proteoglycans may be needed for beaded microfibril formation (see below). Sulphation is a requirement for microfibril assembly, since chlorate treatment ablates microfibril and elastic fiber formation (Robb *et al.*, 1999). This effect may reflect the absence of a proteoglycan, or undersulphation of fibrillins or the microfibril-associated molecule MAGP-1 (see below).

B. *Fibrillin-1 Homotypic Interactions*

The molecular form of fibrillin-1 secreted from cells is unclear. As is the case for most major ECM polymers, fibrillins can undergo limited intracellular assembly to form dimers or trimers that could be intermediates

during extracellular assembly (Ashworth *et al.*, 1999b; Lin *et al.*, 2002; Trask *et al.*, 1999). However, monomers that have been excluded from assembly are often detected in cell culture medium (Wallis *et al.*, 2003; C. M. Kielty, unpublished data). Difficulties in expressing recombinant full-length fibrillin-1 have so far precluded detailed assessment of whether fibrillin-1 is secreted predominantly as a monomer, and whether it can self assemble. Several models of fibrillin-1 alignment within assembled microfibrils have been proposed (see Section V below), and all support the concept that fibrillin-1 can assemble linearly in a head-to-tail fashion at the cell surface. Microfibril assembly must also involve controlled lateral fibrillin-1 packing.

We initially investigated fibrillin-1 C-terminal processing and assembly using an *in vitro* translation system supplemented with semipermeabilized cells as a source of endoplasmic reticulum (Ashworth *et al.*, 1999a). Processing of C-terminal fibrillin-1 was shown to be influenced by N-glycosylation immediately downstream of the furin site, and by association with calreticulin. Interestingly, the C-terminus of fibrillin-2 underwent less efficient processing than C-terminal fibrillin-1 under identical conditions. Differences in processing rates of these two fibrillin isoforms may reflect differential abilities to assemble into microfibrils. Size fractionation of the N-terminal regions of fibrillin-1 (encoded by exons 1–11 and 1–15), and of unprocessed and furin-processed C-terminal fragments of fibrillin-1 (encoded by exons 50–65), revealed that the N-terminus formed abundant disulphide-bonded aggregates. Some association of unprocessed C-terminal fibrillin-1 was also apparent, but processed C-terminal sequences remained monomeric unless N-terminal sequences encoded by exons 12–15 were present. These data indicated the presence of fibrillin-1 molecular recognition sequences within the N-terminus and the extreme C-terminal sequence downstream of the furin site, and a specific N- and C-terminal association that could drive linear accretion of furin-processed fibrillin-1 molecules in the extracellular space.

Other recombinant studies showed that recombinant N-terminal and other regions of fibrillin-1 have a tendency to dimerise (Ashworth *et al.*, 1999b; Trask *et al.*, 1999). An unpaired cysteine residue in the first hybrid domain, which contains nine cysteines, is predicted to be involved in covalently linking N-terminally-aligned fibrillin-1 molecules (Reinhardt *et al.*, 2000b). N- and C-terminal halves of the fibrillin-1 molecule, and the N-terminal half of fibrillin-2, were shown to interact with high affinity in solid-phase and BIAcore binding assays (Tiedemann *et al.*, 2001). However, N- and C-terminal fibrillin-2 polypeptides did not interact with each other. These data demonstrated that fibrillins can directly interact in an N- to C-terminal fashion to form homotypic fibrillin-1 or heterotypic fibrillin-1/fibrillin-2 microfibrils.

We have recently undertaken a detailed binding and kinetic analysis of fibrillin-1 homotypic interactions, using smaller recombinant fragments that span the entire coding region of human fibrillin-1, in order to define more precisely the fibrillin-1 sequences that interact (Marson *et al.*, 2005). The N-terminal sequence encoded by exons 1–8 binds with low nM affinity to itself, and also bound a downstream sequence encoded by exons 9–17. The furin-processed C-terminus and the proteolytically released C-terminal 20 kDa fragment both bind homotypically and tightly to the N-terminal fragment. No other homotypic interactions within fibrillin-1 were detected. MAGP-1, which also binds the N-terminus (see below), was found to inhibit the N- to C-terminal interaction, but not the N- to N-terminal interaction.

In summary, the N- to C-terminal interactions may be a basis for linear fibrillin-1 assembly, while the N- and C-terminal homotypic interactions may be critical in lateral assembly. MAGP-1 may play a role in regulating N- to C-terminal linear fibrillin-1 assembly.

C. Fibrillin-1 Interactions with Microfibril-Associated Molecules

An extensive inventory of microfibril-associated molecules has been identified on the basis of gene mapping of heritable elastic fiber defects, mouse models, and detailed immunohistochemical and biochemical studies of elastic tissues (Kielty *et al.*, 2002a). These molecules can be categorized as those that colocalize or copurify with microfibrils, those that occur at the elastin-microfibril interface or the elastic-fiber-cell interface, and those that are involved in the process of elastic fiber formation. *In vitro* binding studies have confirmed that fibrillin-1 can interact with a number of matrix molecules. They include MAGPs-1 and -2, fibulins-1, -2, and -5, versican, and heparan sulphate. It is now clear that the fibrillin-1 N-terminus is highly interactive. It not only binds itself and the C-terminus, but it can also interact strongly with several of these associated molecules. It is unclear whether these interactions are mutually exclusive.

MAGP-1 (sometimes known as MFAP-2) is possibly the best candidate for an integral microfibril molecule (Gibson *et al.*, 1989; Trask *et al.*, 2000). It is associated with virtually all microfibrils and widely expressed in mesenchymal and connective tissue cells throughout development. Human MAGP-1 is a 183-residue molecule that has two distinctive domains: an acidic N-terminal half that is enriched in proline residues and has a clustering of glutamine residues, and a C-terminal portion that contains 13 cysteine residues and has a net positive charge. MAGP-1 has a matrix-binding domain (MBD) that targets this molecule to the ECM (Segade *et al.*, 2002). MAGP-1 showed calcium-dependent binding to the fibrillin-1 N-terminus (Jensen *et al.*, 2001; Rock *et al.*, 2004). It localizes to

microfibril beads (often two per bead) (Henderson *et al.*, 1996; Kielty and Shuttleworth, 1997) and is probably disulphide bonded to microfibrils since reduction is required for its extraction. MAGP-1 and fibrillin-1 are also both substrates for transglutaminase (Brown-Augsberger *et al.*, 1996; Qian and Glanville, 1997; Rock *et al.*, 2004). MAGP-2, the only other member of this protein family, is a 170–173 residue protein structurally related to MAGP-1, mainly in a central region (Gibson *et al.*, 1998; Segade *et al.*, 2002). MAGP-2 is rich in serine and threonine residues and contains an RGD cell-recognition motif through which it binds to the $\alpha v\beta 3$ integrin (Gibson *et al.*, 1999). MAGP-2 localizes to elastin-associated and elastin-free microfibrils in a number of tissues (Gibson *et al.*, 1998; Hanssen *et al.*, 2004). Its restricted patterns of tissue localization and developmental expression suggest that MAGP-2 is not a primary structural component of microfibrils, but it may be important in cell signaling during microfibril assembly and elastic fibrillogenesis.

Fibrillin-1 interacts with tropoelastin with high affinity, an interaction that is likely to be critical in elastic fiber formation. Using binding assays, we have demonstrated high-affinity calcium-independent binding of two overlapping fibrillin-1 fragments (encoded by central exons 18–25 and 24–30) to tropoelastin, which, in microfibrils, maps to an exposed ''arms'' feature adjacent to the beads (Rock *et al.*, 2004). A further binding site within an adjacent fragment (encoded by exons 9–17) was within an eight-cysteine motif designated TB2 (encoded by exons 16 and 17). A novel transglutaminase crosslink between tropoelastin and fibrillin-1 fragment (encoded by exons 9–17) was localized by mass spectrometry to a sequence encoded by exon 17. The high-affinity binding and crosslinking of tropoelastin to a central fibrillin-1 sequence confirm that this association is fundamental to elastic fiber formation. N-terminally bound MAGP-1 may contribute not only to microfibril assembly, but also to tropoelastin deposition on microfibrils.

Latent TGFβ-binding proteins (LTBPs) comprise repeating cbEGF domains interspersed with TB modules, and are thus members of the ''fibrillin superfamily'' (Oklu and Hesketh, 2000; Sinha *et al.*, 1998). Specific TB modules in LTBP1, LTBP3, and LTBP4 can bind to TGFβ intracellularly, forming a large latent complex that is secreted and cross-linked in the ECM by transglutaminase. Subsequent proteolytic release of the complex is necessary for TGFβ activation, so LTBPs play an important role in tissue targeting of TGFβ. LTBP-1 colocalizes with microfibrils in skin and cell layers of cultured osteoblasts and in embryonic long bone, but not cartilage (Dallas *et al.*, 2000; Raghunath *et al.*, 1998; Taipale *et al.*, 1996). Thus, LTBP-1 is unlikely to be an integral structural component of microfibrils. LTBP-2 colocalizes with fibrillin microfibrils in elastic

fiber-rich tissues (especially in response to arterial injury) and in trabecular bone (Gibson *et al.*, 1995; Kitahama *et al.*, 2000; Sinha *et al.*, 2002). It is a good candidate for an integral microfibril molecule, although it does not bind to TGFβ. The C-terminus of LTBP-3 was reported not to bind fibrillin-1 (Isogai *et al.*, 2003), but it is not known whether LTBP-4 can bind fibrillin-1 (Giltay *et al.*, 1997; Saharinen *et al.*, 1998).

Several proteoglycans are associated with microfibrils and contribute critically to their assembly and integration into the surrounding ECM. Heparin and heparan sulphate bind fibrillin-1 in at least three sites, and supplementing culture medium with heparin inhibits microfibril assembly (Ritty *et al.*, 2003; Tiedemann *et al.*, 2001). Versican, a large chondroitin sulphate proteoglycan of the lectican family, was immunolocalized to microfibrils in skin (Zimmermann *et al.*, 1994). The versican C-terminal lectin domain binds N-terminal fibrillin-1 sequences (Isogai *et al.*, 2002). However, its nonperiodic association with microfibrils indicates that versican is probably not an integral structural component. Instead, it may associate with microfibrils, and its negatively charged chondroitin sulphate chains may influence integration of microfibrils into the surrounding ECM. Polycationic dyes showed an association between proteoglycan and elastic fibers (Baccarani-Contri *et al.*, 1990). Two members of the small leucine-rich PG family, decorin and biglycan, were detected within elastic fibers in dermis; biglycan mapped to the elastin core and decorin mapped to microfibrils. Decorin can interact with both fibrillin-1 and MAGP-1 individually, and together they form a ternary complex (Trask *et al.*, 2000). Small chondroitin sulphate proteoglycans may be associated with microfibrils and contribute to their beaded organization (Kielty *et al.*, 1996).

Fibulin-2 is not crosslinked within microfibrils, but strongly binds a fibrillin-1 N-terminal sequence (within residues 45–450; exons 2–10) in a calcium-dependent interaction (Reinhardt *et al.*, 1996b). Fibulin-1 and fibulin-2 interact with the versican C-terminal lectin domain, and fibulin-2 also binds to aggrecan and brevican (Olin *et al.*, 2000). Fibulin-5 is a critical determinant of elastic fiber formation (see below). It binds strongly to tropoelastin in a calcium-dependent manner although reportedly not to fibrillin-1, and colocalizes with tropoelastin (Nakamura *et al.*, 2002; Yanagisawa *et al.*, 2002).

Microfibril-associated protein-1 (MFAP-1; also known as AMP), MFAP-3, and MFAP-4 (also known as MAGP-36) colocalize with microfibrils and elastic fibers in skin and other tissues (Abrams *et al.*, 1995; Hirano *et al.*, 2002; Horrigan *et al.*, 1992; Lausen *et al.*, 1999; Liu *et al.*, 1997; Toyoshima *et al.*, 1999). In aging and immune conditions, microfibrils can associate with amyloid deposits and accumulate a coating of adhesive glycoproteins, such as vitronectin (Dählback *et al.*, 1990).

D. *Microfibril Maturation*

There is some evidence for extracellular maturation of microfibrils, with the extracellular environment possibly playing a major role in regulating microfibril fate. In human dermal fibroblast cultures, monoclonal antibody 11C1.3 (which binds to beaded microfibrils) does not detect microfibrils until about two weeks in culture, but a polyclonal antibody (PF2) to a fibrillin-1 pepsin fragment can detect abundant microfibrils within three days (Baldock *et al.*, 2001). The time-dependent appearance of 11C1.3-reactive microfibrils suggests some maturation that might be due to conformational changes that unmask cryptic epitopes, transglutaminase crosslinking, or accretion of associated molecules. The relative tissue abundance of heparan sulphate and various microfibril-associated molecules may influence epitope availability and commit microfibrils to distinct extracellular fates.

IV. Microfibril Organization and Elasticity

A. *Organization*

Early transmission electron microscopy (TEM) studies suggested that tissue microfibrils were tubelike structures with a diameter of 10–12 nm (Kielty and Shuttleworth, 1997). The most prominent feature of isolated microfibrils is their ''beads-on-a-string'' appearance with untensioned periodicity of ~56 nm (Keene *et al.*, 1991; Kielty *et al.*, 1991). This is seen after extraction using homogenization, bacterial or purified collagenase with or without hyaluronidase, and by rotary shadowing, negative staining, scanning transmission EM (STEM), atomic force microscopy (AFM), and cryoEM (Baldock *et al.*, 2002; Kielty *et al.*, 2002a,b). Interestingly, by quick-freeze deep etch microscopy, undisturbed zonular microfibrils appear to be more tubular, suggesting that molecular components are lost or that there is a major molecular rearrangement on extraction (Davis *et al.*, 2002). These structural differences between tissue and isolated microfibrils are most prominent in the interbead region.

STEM mass mapping of isolated microfibrils from a variety of developing and adult tissues revealed repeating peaks and troughs of mass (Sherratt *et al.*, 1997). There is abundant evidence for a shoulder of mass to one side of each bead, so microfibrils have an asymmetrical axial mass distribution. There is also evidence for asymmetric mass accretion during development. The axial mass distribution is statistically similar in all repeats, for each given microfibril population, implying that each repeat has comparable molecular composition and organization (Table I).

TABLE I
Fibrillin Microfibril Mass and Periodicity

	Periodicity (nm)	Mass per repeat (kDa)	Bead center mass (kDa/nm)	Interbead center mass (kDa/nm)	References
Fetal*					
Bovine aorta	56.1	2065	62.4	18.9	Sherratt *et al.* (1997)
Bovine nuchal ligament	56.1	2335	60.5	26.9	Sherratt *et al.* (1997)
Adult*					
Canine ciliary zonule	58.9	2550	55.0	28.3	Baldock *et al.* (2001)
Murine skin	55.1	2547	59.8	33.5	Kielty *et al.* (1998)
					Unpublished observations (MJS)
Staggered†					
(N-gly + 8 MAGP-1)					
One-third (domain 18)	54.0	923	19.7	14.6	Lee *et al.* (2004)
One-third (domain 19)	54.0	923	25.2	13.4	Lee *et al.* (2004)
One-third (domain 20)	54.0	923	22.1	13.9	Lee *et al.* (2004)
One-half (domain 28)	81.0	1385	25.0	15.4	—
Hinged†					
Monomer	54.0	2369	62.7	14.7	Baldock *et al.* (2001)
Monomer (N-gly + 8 MAGP-1)	54.0	2772	66.9	18.0	—

*Determined experimentally using STEM mass mapping.

†Determined from theoretical axial mass distributions at a one-third stagger and a one-half stagger (overlap indicated by domain number), and for a hinged model in the presence or absence of N-linked glycosylation (N-gly) and associated MAGP-1.

Calcium chelation and addition experiments have shown that bound calcium profoundly influences microfibril periodicity and organization (Cardy *et al.*, 1998; Kielty and Shuttleworth, 1993; Wess *et al.*, 1998a). In the absence of calcium, microfibril periodicity is reduced and microfibrils have a wider diameter and appear more flexible. We generated three-dimensional reconstructions of isolated microfibrils in the presence of calcium, based on automated electron tomography (Baldock *et al.*, 2001). These data provided fine structural details and supported a fibrillin-1 assembly model in which initial assemblies would undergo conformational maturation to a reversibly extensible beaded polymer (see Section V below), and also suggested a molecular explanation for microfibril extensibility. There was evidence for two arms emerging from one side of the beads, interbead twisting, and the beads appeared more heart-shaped than round. Antibody epitope mapping showed that the N- and C-termini were at either side of the bead, and that there is complex intramolecular folding in 56 nm microfibrils. Electron microscopy images of isolated microfibrils and detailed STEM mass mapping indicate that there are eight molecules in cross-section (Baldock *et al.*, 2001; Wallace *et al.*, 1991). Dehydrated and hydrated isolated microfibrils all show the same major features.

In tissues, microfibrils form loosely packed parallel bundles. X-ray fiber diffraction of hydrated zonular microfibril bundles identified one-third-staggered ''junctions'' that may modulate force transmission (Wess *et al.*, 1998a), and quick-freeze deep-etch analysis of zonules has identified molecular links between microfibrils (Davis *et al.*, 2002).

B. *Elasticity*

Fibrillin-rich microfibrils have endowed tissues with elasticity throughout multicellular evolution. X-ray studies and mechanical testing of microfibril bundles showed that bound calcium influences load deformation, but is not necessary for high extensibility and elasticity (Eriksen *et al.*, 2001; Wess *et al.*, 1997, 1998a,b). Thus, tissue microfibril elasticity is modified by, but not dependent on, calcium-induced beaded periodic changes. This is consistent with the molecular folding model.

We used molecular combing to determine Young's modulus for individual microfibrils, and X-ray diffraction of zonular filaments of the eye to establish the linearity of microfibril periodic extension (Sherratt *et al.*, 2003). Microfibril periodicity is not altered at physiological zonular tissue extensions, and Young's modulus is between 78 to 96 MPa, MPa, which is two orders of magnitude stiffer than elastin. Thus, elasticity

in microfibril-containing tissues arises primarily from reversible alterations in supramicrofibrillar arrangements rather than from intrinsic elastic properties of individual microfibrils, which instead act as reinforcing fibers in fibrous composite tissues.

What can be the molecular explanation for microfibril elasticity? Three-dimensional microfibril reconstructions allowed us to develop a model of fibrillin alignment in extensible microfibrils in which intramolecular folding would act as a molecular ''engine,'' driving extension and recoil (Baldock *et al.*, 2001; Kielty *et al.*, 2002c, 2004). Our model predicts that maturation from initial parallel head-to-tail alignment (~160 nm) to an approximately one-third stagger (~100 nm) occurs by folding at the termini and the proline-rich region, which would align known fibrillin-1 transglutaminase crosslink sequences. Microfibril elasticity (in the range of 56–100 nm) would arise from further intramolecular folding at flexible sites, which could be links between specific TB motifs and cbEGF-like domains. The crystal structure of fibrillin-1 TB4 and flanking cbEGF-like domains has led to another suggested model for microfibril elasticity, based on a staggered arrangement for fibrillin-1 within microfibrils with molecular movement around certain domain linkages (Lee *et al.*, 2004) (see Section V below and Table I).

V. Fibrillin-1 Alignment Models

There are two major current models of fibrillin-1 alignment in microfibrils: the ''molecular hinge'' model of unstaggered molecules and the one-third staggered arrangement. One driver in both models is the presence of an intermolecular fibrillin-1 transglutaminase crosslink (Qian and Glanville, 1997). However, our recent mass spectrometry study has revealed that not every fibrillin-1 molecule is crosslinked within tissue microfibrils (Cain *et al.*, 2005).

A. *Hinged Model*

Our model is based on detailed STEM mass mapping, automated electron tomography, and AFM data (Fig. 2; Baldock *et al.*, 2001, 2002; Kielty *et al.*, 2002c; Sherratt *et al.*, 2003). It predicts maturation from an initial parallel head-to-tail alignment to an approximately one-third stagger (~100 nm) that would facilitate transglutaminase crosslink formation. The staggered form would further pack into a more energetically favorable ~56 nm ''untensioned'' form. The model can account for known microfibril structural features and mass profiles (axial distributions, mass per repeat, mass contrast), is in agreement with the observed number of

molecules in cross-section, and provides a rationale for microfibril elasticity through molecular (but not domain) unfolding. It does not require the presence of additional molecules, but can accommodate associated molecules. For example, tropoelastin could bind at the shoulder in developing

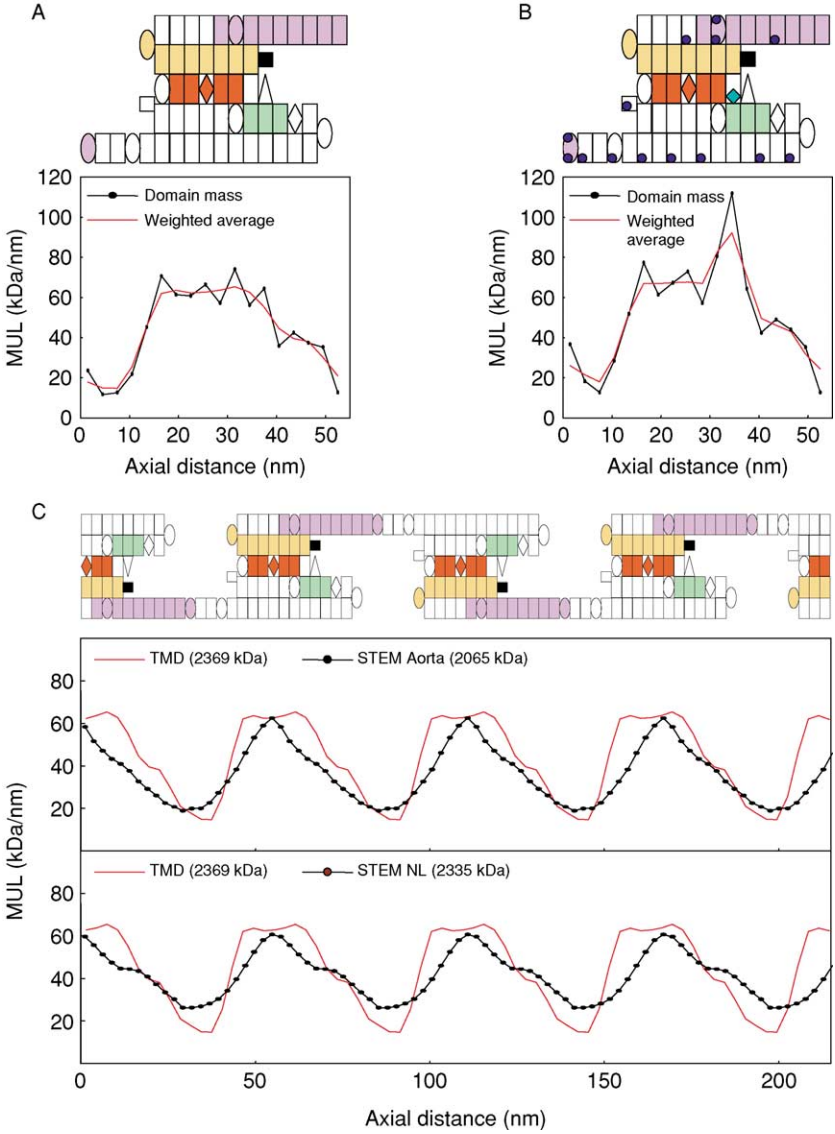


FIG. 2. (*continued*)

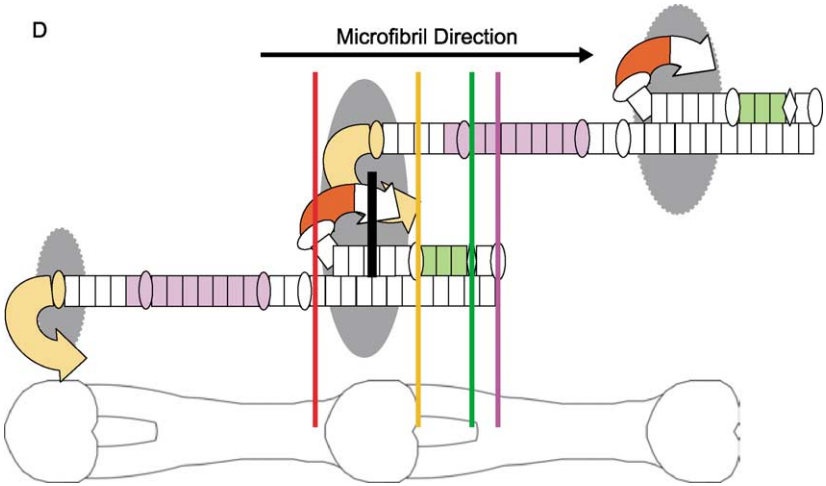


FIG. 2. Theoretical axial mass distributions (TAMD) calculated for the hinge model of fibrillin-1 microfibril structure (Baldock *et al.*, 2001). The mass of each domain was determined from the primary amino acid sequence, with the addition of one Ca^{2+} (40 Da) to each calcium binding EGF domain (cbEGF). All domains were assumed to have an axial extent of 3 nm. (A) TAMD calculated for the primary amino acid sequence + Ca^{2+} of a single repeating unit (bead and interbead). (B) TAMD of a fully glycosylated single repeat (all putative N-linked glycosylation sites occupied with 2.46 kDa carbohydrate chains) and an N-terminal associated MAGP-1 (20.83 kDa). (C) Comparison of theoretical (primary sequence + Ca^{2+}, 8 monomers in cross section) and experimental (fetal bovine aorta and nuchal ligament [NL]) axial mass distributions (Sherratt *et al.*, 1997). (D) Schematic diagram of the intramolecular hinge model described in Baldock *et al.* (2001). The experimentally determined antibody binding positions are indicated on a schematic diagram of a microfibril. The position of a putative transglutaminase crosslink is represented as a black line.

tissues, and two MAGP-1 molecules could bind at each bead (as observed by immunogold EM) (Henderson *et al.*, 1996; Kielty and Shuttleworth, 1997).

The epitopes of the fibrillin-1 monoclonal antibodies, 2502 and 11C1.3, were further characterized as shown in Fig. 1C. A library of overlapping recombinant fibrillin protein fragments (Rock *et al.*, 2004) were screened by Western blot analysis. The epitope of antibody 2502 was found to be within exons 4–8 (amino acids 116–329) and the epitope of antibody 11C1.3 was within exons 18–20 (amino acids 723–909). The position of these exons in the intramolecular hinge model is still consistent with

the experimentally determined location of the antibodies on microfibrils (Fig. 2D).

B. *Staggered Model*

A one-third staggered arrangement has been suggested by extrapolation of molecular dimensions (Downing *et al.*, 1996) and organization from the crystal structure of fibrillin-1 domain arrays (Lee *et al.*, 2004) (Fig. 3). Our calculations show that that this model would require ~24 laterally aligned staggered fibrillin-1 molecules or the additional presence of ~80 MAGP-1 molecules per repeat (for a microfibril with 8 monomers in cross-section) in order to correspond to real microfibril mass maps and mass contrast levels. However, available evidence to date (STEM, rotary shadowing, and quick-freeze-deep etch imaging of sonicated zonular microfibrils) indicates that there are only about eight molecules in a microfibril cross-section (Baldock *et al.*, 2001; Wallace *et al.*, 1991). Elasticity in the staggered model would arise from flexibility at specific TB-cbEGF-like domain junctions. However, in any one molecule, there is a maximum of two TB-cbEGF junctions per interbead repeat, irrespective of the precise region of each staggered molecule within that interbead repeat. Therefore, assuming approximately equal extension of all laterally aligned molecules, that would give only up to ~10 nm extension per repeat. However, the region between the beads, which measures ~35 nm in untensioned microfibrils, has been seen to extend to ~150 nm, so it is difficult to see how such extension could be accounted for on the basis of a maximum of two flexible domain junctions per repeat. It is also not clear what molecular mechanism might stop molecules sliding apart in this staggered model. Crosslinks would be more critical for stabilizing this model.

The refined epitope positions for antibodies 2502 and 11C1.3 were mapped onto the one-third stagger model and compared with the experimentally determined location on microfibrils. There is a discrepancy in the position of the epitope of 11C1.3 in this model. Antibody 11C1.3 binds to the arms/shoulder region of a microfibril at 41.1% of the bead-to-bead distance (Baldock *et al.*, 2001). In the one-third stagger model, exons 18–20 would be located very close to the bead, approximately 15 nm from their actual position (Fig. 3E). If the microfibril direction is reversed, then it is possible to find a better correlation between epitope position and the actual location on a microfibril. However, this model would place the N- and C-termini in the interbead region and the predicted mass profile would have a trough at the bead and a peak in the interbead (Fig. 3F).

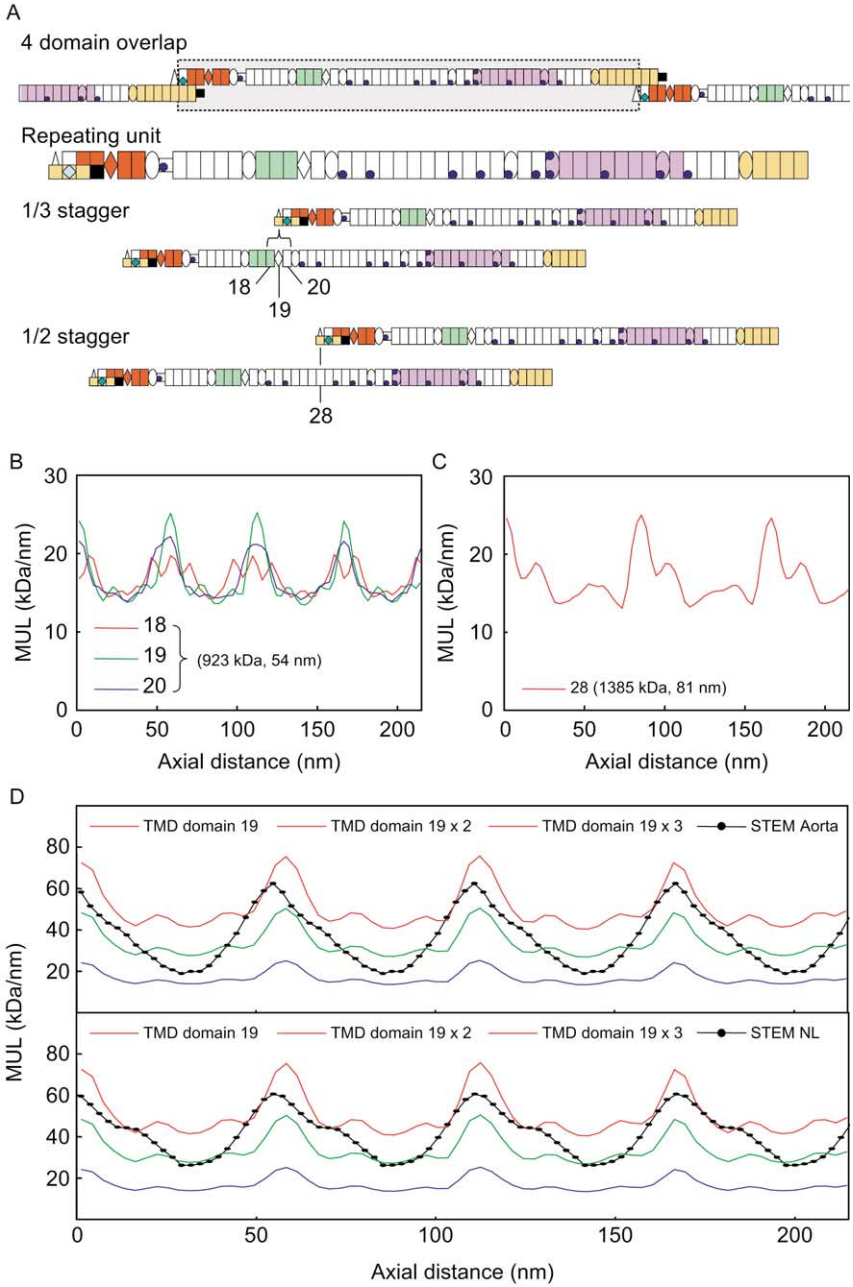


FIG. 3. (*continued*)

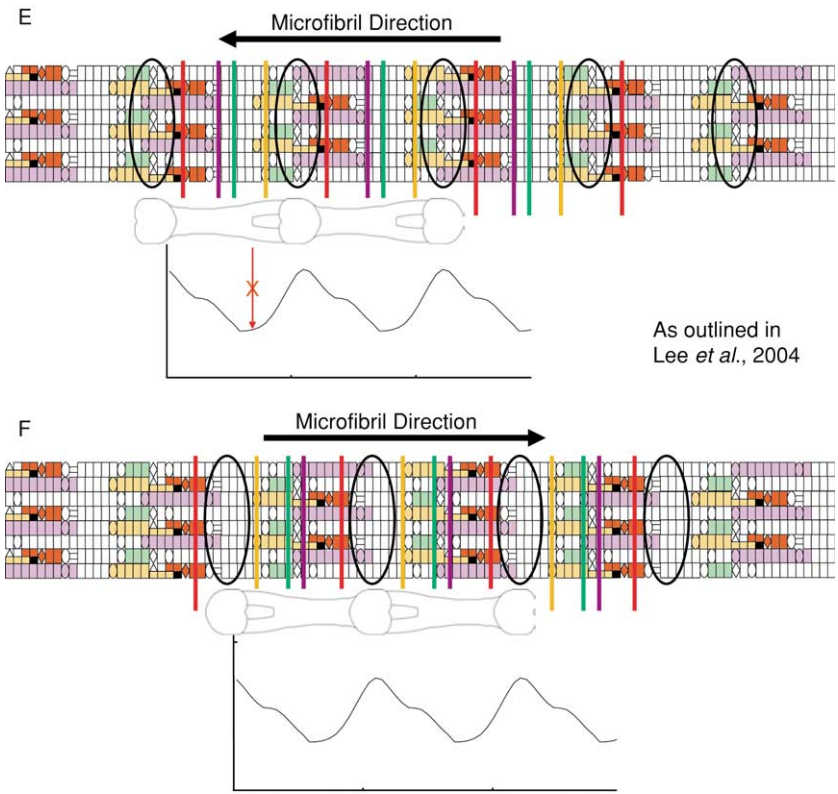


FIG. 3. TAMD calculated for staggered models of fibrillin-1 microfibril structure with 8 monomers in cross section. (A) Lee *et al.* (2004) proposed a one-third staggered model with a four-domain overlap (assuming axial domain dimensions of 3 nm) of the N/C termini. The 55-domain repeating unit can be aligned at one-third and one-half staggers to produce mass peaks with periodicities of 54 and 81 nm, respectively. TAMD (primary sequence + Ca^{2+} + N-linked glycosylations + MAGP-1) calculated for (B) an approximately one-third stagger at domains 18, 19, and 20 and (C) for a 1/2 stagger at domain 28. (D) Comparison of theoretical (domain 19) and experimental (fetal bovine aorta and nuchal ligament) axial mass distributions (Sherratt *et al.*, 1997). Theoretical mass distributions are depicted for 8 (TAMD domain 19), 16 (TAMD domain 19 × 2), or 24 (TAMD domain 19 × 3) monomers in crosssection. The total mass per repeat is 923, 1846, or 2769 kDa for 8, 16, and 24 monomers, respectively. (E) Schematic diagram of the one-third stagger model described in Lee *et al.* (2004). The fibrillin molecule is colored as described above. Experimentally determined antibody epitope positions are indicated and compared with the predicted positions based on this model. There is a discrepancy with the position of antibody 11C1.3, the position in the model is approximately 15 nm away from the actual position on a microfibril. Also, the direction of the STEM mass map profile that was superimposed in Lee *et al.* (2004) was incorrect. It has been shown that the arm region (11C1.3 epitope) is in the shoulder of mass off

VI. Fibrillin Microfibrils in Aging and Disease

A. *Aging*

Elastic fibers are remarkably resilient structural components of tissues, and elastin has a half-life of about 70 years. Nevertheless, normal elastic tissue aging is generally associated with a progressive loss of elasticity (e.g., hardening of the arteries, loss of skin elasticity and wrinkling, and reduced pulmonary capacity). These mechanical changes reflect structural degenerative changes including denudation of the microfibrillar mantle surrounding elastic fibers, and exposure of the elastin core to proteolytic degradation. In zonules, which hold the lens in dynamic suspension, reduced integrity of the zonular filaments based on fibrillin microfibrils leads to lens dislocation, and also contributes to farsightedness in later life. Photodamage to skin results in elastosis (appearance of large amounts of disordered elastin) and loss of the microfibrillar network integrity at the dermal-epidermal junction (Watson *et al.*, 1999). This loss of integrity of the elastic fiber network—from the microfibrillar oxytalin fibers of the papillary dermis, to elaunin fibers associated with small amounts of elastin in the middermis, to thick elastic fibers of the reticular dermis—leads to loss of skin elasticity.

Degenerative structural changes to elastic fibers are often caused by proteolysis through the actions of elastases (including serine proteinases, such as neutrophil elastase) and matrix metalloproteinases (including MMP-2, MMP-9 [gelatinases] and MMP-12 [metalloelastase]) (Ashworth *et al.*, 1999c; Hindson *et al.*, 1999; Kielty *et al.*, 1994), but also by age-related glycation, accumulation of molecules such as amyloid P and fibronectin, and loss of components such as proteoglycans as tissue hydrated status is altered in aging. The proteolytic changes may result in direct loss of microfibrils or elastic fibers. Cell adhesion to microfibrils may also be altered.

B. *Fibrillin Heritable Diseases*

Mutations in fibrillin-1 cause the autosomal dominant disorder Marfan syndrome and related disorders, termed *fibrillinopathies* (Robinson and Booms, 2001). MFS is characterized by life-threatening cardiovascular

---

the bead (Sherratt *et al.*, 2001). (F) Schematic diagram of a one-third stagger model with the microfibril direction reversed, and therefore correct with respect to the antibody positions and STEM mass data. This model has the N- and C-termini in the interbead and therefore would have little mass at the predicted bead and a peak of mass in the interbead instead.

disease and severe skeletal and ocular defects. Progressive aortic root dilatation, with risk of aortic dissection, is the most serious clinical problem for most classic MFS patients. Aortic dissection, rupture, or cardiac failure due to mitral/aortic valve regurgitation cause death in over 90% of these patients. Neonatal MFS patients die perinatally of congestive heart failure and cardiac valve insufficiency. Other overlapping fibrillin-1 related pathologies include MASS syndrome (mitral valve prolapse, aortic dilatation, and skin and skeletal manifestations syndrome), and autosomal thoracic aortic aneurysms and dissections (TAAD) (Hasham *et al.*, 2003).

1. Fibrillin-1 Mutations

Over 350 MFS mutations have been identified (Collod-Beroud *et al.*, 2003). Missense mutations represent approximately two-thirds of all fibrillin-1 mutations, the majority affecting cbEGF domains. Mutations that change a cbEGF cysteine or calcium-binding consensus residue usually cause classic MFS, although phenotypic effects are variable and depend on location, calcium-binding characteristics, and domain context. Mutations affecting residues other than cysteines and calcium-binding residues are rarer, and include those that create a cysteine or N-glycosylation site, those that affect a conserved glycine involved in cbEGF domain-domain packing, and mutations of nonconserved residues. Premature truncation codon (PTC) mutations represent ~20% of FBN1 mutations. mRNAs carrying a PTC due to a frameshift or nonsense mutation are generally at reduced levels due to nonsense-mediated decay. Some PTCs cause mild disease, but others cause classic MFS. Splice site mutations represent ~12% of all mutations, are caused by exon skipping and activation of cryptic splice sites, and lead to inframe exon deletions and internally deleted fibrillin-1 molecules.

2. Genotype-Phenotype Relationships

MFS has, at least in part, a dominant negative pathogenesis with mutant fibrillin-1 interfering with wild-type fibrillin-1 secretion, assembly, and/or function. Reduced levels of functional extracellular fibrillin-1 microfibrils impair vascular homeostasis. Genotype-phenotype relationships are poorly understood at the molecular level. Mutations occur throughout fibrillin-1 with some correlations between location and disease severity, reflecting sequence-specific roles in secretion, assembly, and extracellular function. Localized mutation-specific effects also influence phenotype. Exons 24–32 contain the mutations causing neonatal MFS and atypical severe MFS (suffer aortic dissection younger than 16 years of age) but also some mutations causing classic MFS. Mutation location is critical since similar

mutations in different domains can have profoundly different phenotypic outcomes, presumably reflecting different roles for specific sequences in microfibril assembly, structure, and function. Mutation-specific domain misfolding effects can also strongly influence phenotypic outcome.

a. Dominant Negative Pathogenesis In MFS, the products of both wild-type and mutant fibrillin-1 alleles are coexpressed by cells, so mutant fibrillin-1 molecules can interfere with the secretion, assembly, and/or extracellular function of wild-type fibrillin-1. PTC mutations have reduced amounts of mutant transcript and milder phenotypes, supporting the hypothesis that, below a certain threshold of wild-type: mutant fibrillin-1 (mutant *FBN1*, 6–16%), mutant molecules generally do not disrupt microfibrils to such an extent that severe disease results. However, a recent study using transgenesis has revealed that haploinsufficiency for wild-type fibrillin-1, rather than production of mutant protein, could be the primary determinant of failed microfibrillar assembly (Judge *et al.*, 2004).

b. Impaired Tissue Homeostasis *Fbn-1* gene-targeted mice defects imply that fibrillin-1 contributes vitally to vascular homeostasis. The collagen-rich adventitia sustains the bulk of hemodynamic stress, so aortic dilatation in MFS could reflect loss of adventitial tensile strength, possibly due to a direct (undefined) role for microfibrils in the adventitia or by altered medial elastic fibers affecting adventitial function.

c. Increased Susceptibility to Proteolysis When calcium is bound, contiguous cbEGF-like domain arrays adopt a stable rodlike form and microfibrils exhibit 56 nm periodicity and well-defined organization. Bound calcium can protect fibrillin-1 from proteolysis at cryptic proteolytic cleavage sites. We identified a disease-causing mutation (E2447K) that disrupts calcium binding and exposes a cleavage site for pathologically relevant enzymes MMP-9 and MMP-12 (metalloelastase) (Ashworth *et al.*, 1999c). Others have monitored mutation-induced domain conformational changes by assessing exposure of cryptic trypsin cleavage sites.

*d. Role in Regulating TGF*β *Activity* Fibrillin-1 and microfibrils contribute to the regulated activation of members of the TGFβ superfamily of cytokines (Charbonneau *et al.*, 2004). TGFβs are cytokines that regulate many cellular processes including proliferation, cell cycle arrest, apoptosis, differentiation, and ECM formation, through a heterotrimeric complex of type I and type II receptors with serine-threonine kinase activities in their cytoplasmic domains. Selected manifestations of MFS, including progressive pulmonary disease, manifest excessive cytokine activation and signaling

(Neptune *et al.*, 2003). *Fbn-1*-deficient mice have excessive TGFβ activity that probably underlies their tendency to develop emphysema and could explain other manifestations of MFS. Thus, perturbation of TGFβ signaling contributes to the pathogenesis of ECM disorders, but by a mechanism different from that reported in Fbn-1 deficient mice. Heterozygous TGFBR2 mutations can cause a second type of Marfan syndrome (MFS2; Mizuguchi *et al.*, 2004). Four missense mutations have been identified in TGFBR2 in four unrelated probands, which lead to loss of TGFβ signaling activity on ECM formation.

VII. Summary

In summary, fibrillin-1 microfibrils are critically important structural and elastic polymers of the ECM, and are also key regulators of TGFβ activity in development. Recent molecular studies have provided new understanding of how they assemble in the pericellular environment, and testing of molecular alignment models is highlighting how their molecular packing arrangement endows them with their unique elastic properties. This information is contributing to our understanding of fibrillin-1 microfibril function in health and disease.

Acknowledgments

Studies from our laboratory were funded by the Medical Research Council (UK), the Wellcome Trust, the European Union, and the Royal Society.

References

Abrams, W. R., Ma, R. I., Kucich, U., Bashir, M. M., Decker, S., Tsipouras, P., McPherson, J. D., Wasmuth, J. J., and Rosenbloom, J. (1995). Molecular cloning of the microfibrillar protein MFAP3 and assignment of the gene to human-chromosome 5q32-q33.2. *Genomics* **26,** 47–54.

Ashworth, J. L., Kelly, V., Rock, M. J., Shuttleworth, C. A., and Kielty, C. M. (1999a). Regulation of fibrillin carboxy-terminal furin processing by N-glycosylation, and association of amino- and carboxy-terminal sequences. *J. Cell Sci.* **112,** 4163–4171.

Ashworth, J. L., Kelly, V., Wilson, R., Shuttleworth, C. A., and Kielty, C. M. (1999b). Fibrillin assembly: Dimer formation mediated by amino-terminal sequences. *J. Cell Sci.* **112,** 3549–3558.

Ashworth, J. L., Sherratt, M. J., Rock, M. J., Murphy, G., Shapiro, S. D., Shuttleworth, C. A., and Kielty, C. M. (1999c). Fibrillin turnover by metalloproteinases: Implications for connective tissue remodelling. *Biochem. J.* **340,** 171–181.

Ashworth, J. L., Kielty, C. M., and McLeod, D. (2000). Fibrillin and the eye. *Br. J. Ophthalmol.* **84,** 1312–1317.

Baccarani-Contri, M., Vincenzi, D., Cicchetti, F., Mori, G., and Pasquali-Ronchetti, I. (1990). Immunocytochemical localization of proteoglycans within normal elastic fibres. *Eur. J. Cell Biol.* **53,** 305–312.

Baldock, C., Koster, A. J., Ziese, U., Sherratt, M. J., Kadler, K. E., Shuttleworth, C. A., and Kielty, C. M. (2001). The supramolecular organisation of fibrillin-rich microfibrils. *J. Cell Biol.* **152,** 1045–1056.

Baldock, C., Gilpin, C., Koster, A., Ziese, U., Kadler, K. E., Kielty, C. M., and Holmes, D. F. (2002). Three-dimensional reconstructions of extracellular matrix polymers using automated electron tomography. *J. Struct. Biol.* **138,** 130–135.

Bax, D. V., Bernard, S. E., Lomas, A., Morgan, A., Shuttleworth, C. A., Humphries, M. J., and Kielty, C. M. (2003). Cell adhesion to fibrillin-1 molecules and microfibrils is mediated by $\alpha 5\beta 1$ and $\alpha v\beta 3$ integrins. *J. Biol. Chem.* **278,** 34605–34616.

Booms, P., Tiecke, F., Rosenberg, T., Hagemeier, C., and Robinson, P. N. (2000). Differential effect of FBN1 mutations on *in vitro* proteolysis of recombinant fibrillin-1 fragments. *Hum. Genet.* **107,** 216–224.

Brown-Augsberger, P., Broekelmann, T., Mecham, L., Mercer, R., Gibson, M. A., Cleary, E. G., Abrams, W. R., Rosenbloom, J., and Mecham, R. P. (1996). Microfibril-associated glycoprotein binds to the carboxyl-terminal domain of tropoelastin and is a substrate for transglutaminase. *J. Biol. Chem.* **269,** 284423–284449.

Cain, S. A., Morgan, A., Sherratt, M. J., Ball, S. G., Shuttleworth, C. A., and Kielty, C. M. (2005). Proteomic analysis of fibrillin-rich microfibrils. *Proteomics* (in review).

Cardy, C. M., and Handford, P. A. (1998). Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium-binding epidermal growth factor-like domains. *J. Mol. Biol.* **276,** 855–860.

Charbonneau, N. L., Ono, R. N., Corson, G. M., Keene, D. R., and Sakai, L. Y. (2004). Fine tuning of growth factor signals depends on fibrillin microfibril networks. *Birth Defects Res. Part C Embryo Today* **72,** 37–50.

Collod-Beroud, G., Le Bourdelles, S., Ades, L., Ala-Kokko, L., Booms, P., Boxer, M., Child, A., Comeglio, P., De Paepe, A., Hyland, J. C., Holman, K., Kaitila, I., Loeys, B., Matyas, G., Nuytinck, L., Peltonen, L., Rantamaki, T., Robinson, P., Steinmann, B., Junien, C., Beroud, C., and Boileau, C. (2003). Update of the UMD-FBN1 mutation database and creation of an FBN1 polymorphism database. *Hum. Mutat.* **22,** 199–208.

Corson, G. M., Charbonneau, N. L., Keene, D. R., and Sakai, L. Y. (2004). Differential expression of fibrillin-3 adds to microfibril variety in human and avian, but not rodent, connective tissues. *Genomics* **83,** 461–472.

Dahlbäck, K., Ljungquist, A., Lofberg, H., Dahlbäck, B., Engvall, E., and Sakai, L. Y. (1990). Fibrillin immunoreactive fibers constitute a unique network in the human dermis: Immunohistochemical comparison of the distributions of fibrillin, vitronectin, amyloid P component, and orcein stainable structures in normal skin and elastosis. *J. Invest. Dermatol.* **94,** 284–291.

Dallas, S. L., Keene, D. R., Saharinen, J., Sakai, L. Y., Mundy, G. R., and Bonewald, L. F. (2000). Role of the latent transforming growth factor beta binding protein family in fibrillin-containing microfibrils in bone cells *in vitro* and *in vivo*. *J. Bone Mineral Res.* **15,** 68–81.

Davis, E. C., Roth, R. A., Heuser, J. E., and Mecham, R. P. (2002). Ultrastructural properties of ciliary zonule microfibrils. *J. Struct. Biol.* **139,** 65–75.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D., and Handford, P. A. (1996). Solution structure of a pair of calcium-binding epidermal growth factor-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell* **85,** 597–605.

Eriksen, T. A., Wright, D. M., Purslow, P. P., and Duance, V. C. (2001). Role of Ca^{2+} for the mechanical properties of fibrillin. *Proteins: Str. Funct. Genet.* **45,** 90–95.

Gibson, M. A., Kumaratilake, J. S., and Cleary, E. G. (1989). The protein components of the 12-nanometer microfibrils of elastic and non-elastic tissues. *J. Biol. Chem.* **264,** 4590–4598.

Gibson, M. A., Hatzinikolas, G., Davis, E. C., Baker, E., Sutherland, G. R., and Mecham, R. P. (1995). Bovine latent transforming growth factor β1-binding protein-2: Molecular cloning, identification of tissue isoforms, and immunolocalization to elastin-associated microfibrils. *Mol. Cell. Biol.* **15,** 6932–6942.

Gibson, M. A., Finnis, M. L., Kumaratilake, J. L., and Cleary, E. G. (1998). Microfibril-associated glycoprotein-2 (MAGP-2) is specifically associated with fibrillin-containing microfibrils and exhibits more restricted patterns of tissue localisation and developmental expression than its structural relative MAGP-1. *J. Histochem. Cytochem.* **46,** 871–885.

Gibson, M. A., Leavesley, D. I., and Ashman, L. K. (1999). Microfibril-associated glycoprotein-2 specifically interacts with a range of bovine and human cell types via alphaVbeta3 integrin. *J. Biol. Chem.* **274,** 13060–13065.

Giltay, R., Kostka, G., and Timpl, R. (1997). Sequence and expression of a novel member (LTBP-4) of the family of latent transforming growth factor-β binding proteins. *FEBS Lett.* **411,** 164–168.

Handford, P. A., Downing, A. K., Reinhardt, D. P., and Sakai, L. Y. (2000). Fibrillin: From domain structure to supramolecular assembly. *Matrix Biol.* **19,** 457–470.

Hanssen, E., Hew, F. H., Moore, E., and Gibson, M. A. (2004). MAGP-2 has multiple binding regions on fibrillins and has covalent periodic association with fibrillin-containing microfibrils. *J. Biol. Chem.* **279,** 29185–29194.

Hasham, S. N., Willing, M. C., Guo, D. C., Muilenburg, A., He, R., Tran, V. T., Scherer, S. E., Shete, S. S., and Milewicz, D. M. (2003). Mapping a locus for familial thoracic aortic aneurysms and dissections (TAAD2) to 3p24–25. *Circulation* **107,** 3184–3190.

Henderson, M., Polewski, R., Fanning, J. C., and Gibson, M. A. (1996). Microfibril-associated glycoprotein-1 (MAGP-1) is specifically located on the beads of the beaded-filament structure for fibrillin-containing microfibrils as visualized by the rotary shadowing technique. *J. Histochem. Cytochem.* **44,** 1389–1397.

Hindson, V. J., Ashworth, J. L., Rock, M. J., Cunliffe, S., Shuttleworth, C. A., and Kielty, C. M. (1999). Fibrillin degradation by matrix metalloproteinases: Identification of amino- and carboxy-terminal cleavage sites. *FEBS Lett.* **452,** 195–198.

Hirano, E., Fujimoto, N., Tajima, S., Akiyama, M., Ishibashi, R., and Okamoto, K. (2002). Expression of 36-kDa microfibril-associated glycoprotein (MAGP-36) in human keratinocytes and its localization in skin. *J. Dermatol. Sci.* **28,** 60–67.

Horrigan, S. K., Rich, C. B., Streeten, B. W., Li, Z. Y., and Foster, J. A. (1992). Characterization of an associated microfibril protein through recombinant DNA techniques. *J. Biol. Chem.* **267,** 10087–10095.

Isogai, Z, Aspberg, A., Keene, D. R., Ono, R. N., Reinhardt, D. P., and Sakai, L. Y. (2002). Versican interacts with fibrillin-1 and links extracellular microfibrils to other connective tissue networks. *J. Biol. Chem.* **277,** 4565–4572.

Isogai, Z., Ono, R. N., Ushiro, S., Keene, D. R., Chen, Y., Mazzieri, R., Charbonneau, N. L., Reinhardt, D. P., Rifkin, D. B., and Sakai, L. Y. (2003). Latent transforming growth factor beta-binding protein 1 interacts with fibrillin and is a microfibril-associated protein. *J. Biol. Chem.* **278,** 2750–2757.

Jensen, S. A., Reinhardt, D. P., Gibson, M. A., and Weiss, A. S. (2001). Protein interaction studies of MAGP-1 with tropoelastin and fibrillin-1. *J. Biol. Chem.* **276,** 39661–39666.

Judge, D. P., Biery, N. J., Keene, D. R., Geubtner, J., Myers, L., Huso, D. L., Sakai, L. Y., and Dietz, H. C. (2004). Evidence for a critical contribution of haploinsufficiency in the complex pathogenesis of Marfan syndrome. *J. Clin. Invest.* **114,** 172–181.

Keene, D. R., Maddox, B. K., Kuo, H-J., Sakai, L. Y., and Glanville, R. W. (1991). Extraction of extensible beaded structure and their identification as extracellular matrix fibrillin-containing microfibrils. *J. Histochem. Cytochem.* **39,** 441–449.

Kielty, C. M., Cummings, C., Whittaker, S. P., Shuttleworth, C. A., and Grant, M. E. (1991). Isolation and ultrastructural analysis of intact high-M_r assemblies of type VI collagen and fibrillin. *J. Cell Sci.* **99,** 797–807.

Kielty, C. M., and Shuttleworth, C. A. (1993). The role of calcium in the organisation of fibrillin microfibrils. *FEBS Lett.* **336,** 323–326.

Kielty, C. M., Woolley, D. E., Whittaker, S. P., and Shuttleworth, C. A. (1994). Catabolism of intact fibrillin microfibrils by neutrophil elastase, chymotrypsin and trypsin. *FEBS Lett.* **351,** 85–89.

Kielty, C. M., Whittaker, S. P., and Shuttleworth, C. A. (1996). Fibrillin: Evidence that chondroitin sulphate proteoglycans are components of microfibrils and associate with newly synthesised monomers. *FEBS Lett.* **386,** 169–173.

Kielty, C. M., and Shuttleworth, C. A. (1997). Microfibrillar elements of the dermal matrix. *Microsc. Res. Tech.* **38,** 407–427.

Kielty, C. M., Raghunath, M., Siracusa, L. D., Sherratt, M. J., Peters, R., Shuttleworth, C. A., and Jimenez, S. A. (1998). The Tight skin mouse: Demonstration of mutant fibrillin-1 production and assembly into abnormal microfibrils. *J. Cell Biol.* **140,** 1159–1166.

Kielty, C. M., Sherratt, M. J., Haston, J. L., Wess, T. J., and Shuttleworth, C. A. (2002a). Fibrillin-rich microfibrils: Elastic biopolymers of the extracellular matrix. *J. Muscle Res. Cell Motil.* **23,** 581–596.

Kielty, C. M., Sherratt, M. J., and Shuttleworth, C. A. (2002b). Elastic fibres. *J. Cell Sci.* **115,** 2817–2828.

Kielty, C. M., Baldock, C., Rock, M. J., Ashworth, J. L., and Shuttleworth, C. A. (2002c). Fibrillin: From microfibril assembly to biomechanical function. *Phil. Trans. R. Soc. Lond. B.* **357,** 207–217.

Kielty, C. M., Wess, T. J., Haston, J. L., Sherratt, M. J., Baldock, C., and Shuttleworth, C. A. (2004). Organization and biomechanical properties of fibrillin microfibrils. *In* ''Marfan Syndrome'' (P. Robinson and M. Godfrey, Eds.), Landes Bioscience/Eurekah.com.

Kitahama, S., Gibson, M. A., Hatzinikolas, G., Hay, S., Kuliwaba, J. L., Evdokiou, A., Atkins, G. J., and Findlay, D. M. (2000). Expression of fibrillins and other microfibril-associated proteins in human bone and osteoblast-like cells. *Bone* **27,** 61–67.

Lack, J., O'Leary, J. M., Knott, V., Yuan, X., Rifkin, D. B., Handford, P. A., and Downing, A. K. (2003). Solution structure of the third TB domain from LTBP1 provides insight into assembly of the large latent complex that sequesters latent TGF-beta. *J. Mol. Biol.* **334,** 281–291.

Lausen, M., Lynch, N., Schlosser, A., Tornoe, I., Sackmose, S. G., Teisner, B., Willis, A. C., Crouch, E., Schwaeble, W., and Holmskov, U. (1999). Microfibril-associated protein 4 is present in lung washings and binds to the collagen region of lung surfactant protein D. *J. Biol. Chem.* **274,** 32234–32240.

Lee, S. S., Knott, V., Jovanovic, J., Harlos, K., Grimes, J. M., Choulier, L., Mardon, H. J., Stuart, D. I., and Handford, P. A. (2004). Structure of the integrin binding fragment from fibrillin-1 gives new insights into microfibril organization. *Structure* **12,** 717–729.

Lin, G., Tiedemann, K., Vollbrandt, T., Peters, H., Batge, B., Brinckmann, J., and Reinhardt, D. P. (2002). Homo- and heterotypic fibrillin-1 and -2 interactions constitute the basis for the assembly of microfibrils. *J. Biol. Chem.* **277,** 50795–50804.

Liu, W. G., Faraco, J., Qian, C. P., and Francke, U. (1997). The gene for microfibril-associated protein-1 (MFAP1) is located several megabases centromeric to FBN1 and is not mutated in Marfan syndrome. *Human Genet.* **99,** 578–584.

Marson, A., Rock, M. J., Cain, S. A., Freeman, L., Morgan, A., Mellody, K., Sherratt, M. J., Shuttleworth, C. A., Baldock, C., and Kielty, C. M. (2005). Homotypic fibrillin-1 interactions in microfibril assembly. *J. Biol. Chem.* (in press).

Mecham, R. P., and Davis, E. C. (1994). Elastic fiber structure and assembly. *In* ''Extracellular Matrix Assembly and Structure'' (P. D. Yurchenco, D. E. Birk, and R. P. Mecham, Eds.), pp. 281–314. Academic Press, New York.

Mizuguchi, T., Collod-Beroud, G., Akiyama, T., Abifadel, M., Harada, N., Morisaki, T., Allard, D., Varret, M., Claustres, M., Morisaki, H., Ihara, M., Kinoshita, A., Yoshiura, K., Junien, C., Kajii, T., Jondeau, G., Ohta, T., Kishino, T., Furukawa, Y., Nakamura, Y., Niikawa, N., Boileau, C., and Matsumoto, N. (2004). Heterozygous TGFBR2 mutations in Marfan syndrome. *Nat. Genet.* **36,** 855–860.

Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R., and Ohara, O. (2001). Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain with code for large proteins *in vitro*. *DNA Res.* **8,** 85–95.

Nakamura, T., Lozano, P. R., Ikeda, Y., Iwanaga, Y., Hinek, A., Minamisawa, S., Cheng, C-F., Kobuke, K., Dalton, N., Takada, Y., Tashiro, K., Ross, J., Honjo, T., and Chien, K. R. (2002). Fibulin-5/DANCE is essential for elastogenesis *in vivo*. *Nature* **415,** 171–175.

Neptune, E. R., Frischmeyer, P. A., Arking, D. E., Myers, L., Bunton, T. E., Gayraud, B., Ramirez, F., Sakai, L. Y., and Dietz, H. C. (2003). Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nat. Genet.* **33,** 407–411.

Oklu, R., and Hesketh, R. (2000). The latent transforming growth factor beta binding protein (LTBP) family. *Biochem. J.* **352,** 601–610.

Olin, A. I., Morgelin, M., Sasaki, T., Timpl, R., Heinegärd, D., and Aspberg, A. (2000). The proteoglycans aggrecan and versican form networks with fibulin-2 through their lectin domain binding. *J. Biol. Chem.* **276,** 1253–1261.

Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J. R., Sykes, B., Pangilinan, T., and Bonadio, J. (1993). Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum. Mol. Genet.* **2,** 961–968.

Pfaff, M., Reinhardt, D. P., Sakai, L. Y., and Timpl, R. (1996). Cell adhesion and integrin binding to recombinant human fibrillin-1. *FEBS Lett.* **384,** 247–250.

Qian, R. Q., and Glanville, R. W. (1997). Alignment of fibrillin molecules in elastic microfibrils is defined by transglutaminase-derived cross-links. *Biochemistry* **36,** 15841–15847.

Raghunath, M., Unsold, C., Kubitscheck, U., Bruckner-Tuderman, L., Peters, R., and Meuli, M. (1998). The cutaneous microfibrillar apparatus contains latent transforming

growth factor-beta binding protein-1 (LTBP-1) and is a repository for latent TGF-beta 1. *J. Invest. Dermatol.* **111,** 559–564.

Raghunath, M., Putnam, E. A., Ritty, T., Hamstra, D., Park, E. S., Tschodrich-Rotter, M., Peters, R., Rehemtulla, A., and Milewicz, D. M. (1999). Carboxy-terminal conversion of profibrillin to fibrillin at a basic site by PACE/furin-like activity required for incorporation in the matrix. *J. Cell Sci.* **112,** 1093–1100.

Reinhardt, D. P., Keene, D. R., Corson, G. M., Poschl, E., Bachinger, H. P., Gambee, J. E., and Sakai, L. Y. (1996a). Fibrillin-1: Organization in microfibrils and structural properties. *J. Mol. Biol.* **258,** 104–116.

Reinhardt, D. P., Sasaki, T., Dzamba, B. J., Keene, D. R., Chu, M-L., Göhring, W., Timpl, R., and Sakai, L. Y. (1996b). Fibrillin-1 and fibulin-2 interact and are colocalized in some tissues. *J. Biol. Chem.* **271,** 19489–19496.

Reinhardt, D. P., Mechling, D. E., Boswell, B. A., Keene, D. R., Sakai, L. Y., and Bachinger, H. P. (1997). Calcium determines the shape of fibrillin. *J. Biol. Chem.* **272,** 7368–7373.

Reinhardt, D. P., Ono, R. N., Notbohm, H., Muller, P. K., Bachinger, H. P., and Sakai, L. Y. (2000a). Mutations in calcium-binding epidermal growth factor modules render fibrillin-1 susceptible to proteolysis. A potential disease-causing mechanism in Marfan syndrome. *J. Biol. Chem.* **275,** 12339–12345.

Reinhardt, D. P., Gambee, J. E., Ono, R. N., Bachinger, H. P., and Sakai, L. Y. (2000b). Initial steps in assembly of microfibrils. Formation of disulfide-cross-linked multimers containing fibrillin-1. *J. Biol. Chem.* **275,** 2205–2210.

Ritty, T. M., Broekelmann, T., Tisdale, C., Milewicz, D. M., and Mecham, R. P. (1999). Processing of the fibrillin-1 carboxyl-terminal domain. *J. Biol. Chem.* **274,** 8933–8940.

Ritty, T. M., Broekelmann, T. J., Werneck, C. C., and Mecham, R. P. (2003). Fibrillin-1 and -2 contain heparin-binding sites important for matrix deposition and that support cell attachment. *Biochem. J.* **375,** 425–432.

Robinson, P. N., and Booms, P. (2001). The molecular pathogenesis of the Marfan syndrome. *Cell Mol. Life Sci.* **58,** 1698–1707.

Robb, B. W., Wachi, H., Schaub, T., Mecham, R. P., and Davis, E. C. (1999). Characterization of an *in vitro* model of elastic fiber assembly. *Mol. Biol. Cell* **10,** 3595–3605.

Rock, M. J., Cain, S., Freeman, L., Morgan, A., Mellody, K., Marson, A., Shuttleworth, C. A., Weiss, A. S., and Kielty, C. M. (2004). Molecular basis of elastic fiber formation: Critical interactions and a tropoelastin – fibrillin-1-crosslink. *J. Biol. Chem.* **279,** 23748–23758.

Saharinen, J., Taipale, J., Monni, O., and Keski-Oja, J. (1998). Identification and characterization of a new latent transforming growth factor-β-binding protein, LTBP-4. *J. Biol. Chem.* **273,** 18459–18469.

Sakamoto, H., Broekelmann, T., Cheresh, D. A., Ramirez, F., Rosenbloom, J., and Mecham, R. P. (1996). Cell-type specific recognition of RGD- and non-RGD-containing cell binding domains in fibrillin-1. *J. Biol. Chem.* **271,** 4916–4922.

Sechler, J. L., Rao, H., Cumiskey, A. M., Vega-Colón, I., Smith, M. S., Murata, T., and Schwarzbauer, J. E. (2001). A novel fibronectin binding site required for fibronectin fibril growth during matrix assembly. *J. Cell Biol.* **154,** 1081–1088.

Segade, F., Trask, B. C., Broekelmann, T. J., Pierce, R. A., and Mecham, R. P. (2002). Identification of a matrix binding domain in the microfibril-associated glycoproteins 1 and 2 (MAGP1 and 2) and intracellular localization of alternative splice forms. *J. Biol. Chem.* **277,** 11050–11057.

Sherratt, M. J., Holmes, D. F., Shuttleworth, C. A., and Kielty, C. M. (1997). Scanning transmission electron microscopy mass analysis of fibrillin-containing microfibrils from foetal elastic tissues. *Int. J. Biochem. Cell Biol.* **29,** 1063–1070.

Sherratt, M. J., Wess, T. J., Baldock, C., Ashworth, J., Purslow, P. P., Shuttleworth, C. A., and Kielty, C. M. (2001). Fibrillin-rich microfibrils of the extracellular matrix: Ultrastructure and assembly. *Micron* **32,** 185–200.

Sherratt, M. J., Holmes, D. F., Haston, J. L., Wess, T. J., Shuttleworth, C. A., and Kielty, C. M. (2003). Fibrillin microfibrils are stiff reinforcing fibres in compliant tissues. *J. Mol. Biol.* **332,** 183–193.

Sherratt, M. J., Holmes, D. F., Shuttleworth, C. A., and Kielty, C. M. (2004). Substrate dependent morphology of supra-molecular assemblies: Fibrillin and type VI collagen microfibrils. *Biophys. J.* **86,** 3211–3222.

Sinha, S., Nevett, C., Shuttleworth, C. A., and Kielty, C. M. (1998). Cellular and extracellular biology of the latent transforming growth factor-beta binding proteins. *Matrix Biol.* **17,** 529–545.

Sinha, S., Shuttleworth, C. A., Heagerty, A. M., and Kielty, C. M. (2002). Expression of latent TGFbeta binding proteins and association with TGFbeta-1 and fibrillin-1 in the response to arterial injury. *Cardiovasc. Res.* **53,** 971–983.

Smallridge, R. S., Whiteman, P., Werner, J. M., Campbell, I. D., Handford, P. A., and Downing, A. K. (2003). Solution structure and dynamics of a calcium binding epidermal growth factor-like domain pair from the neonatal region of human fibrillin-1. *J. Biol. Chem.* **278,** 12199–12206.

Taipale, J., Saharinen, J., Hedman, K., and Keski-Oja, J. (1996). Latent transforming growth-factor-β1 and its binding-protein are components of extracellular-matrix microfibrils. *J. Histochem. Cytochem.* **44,** 875–889.

Tiedemann, K., Bätge, B., Müller, P. K., and Reinhardt, D. P. (2001). Interactions of fibrillin-1 with heparin/heparan sulphate, implications for microfibrillar assembly. *J. Biol. Chem.* **276,** 36035–36042.

Toyoshima, T., Yamashita, K., Shishibori, T., Itano, T., and Kobayashi, R. (1999). Ultrastructural distribution of 36-kDa microfibril-associated glycoprotein (MAGP-36) in human and bovine tissues. *J. Histochem. Cytochem.* **47,** 1049–1056.

Trask, T. M., Ritty, T. M., Broekelmann, T., Tisdale, C., and Mecham, R. P. (1999). N-terminal domains of fibrillin 1 and fibrillin 2 direct the formation of homodimers: A possible first step in microfibril assembly. *Biochem J.* **340,** 693–701.

Trask, B. C., Trask, T. M., Broekelmann, T., and Mecham, R. P. (2000). The microfibrillar proteins MAGP-1 and fibrillin-1 form a ternary complex with the chondroitin sulfate proteoglycan decorin. *Mol. Biol. Cell* **11,** 1499–1507.

Wallace, R. N., Streeten, B. W., and Hanna, R. B. (1991)). Rotary shadowing of elastic system microfibrils in the ocular zonule, vitreous, and ligamentum nuchae. *Curr. Eye Res.* **10,** 99–109.

Wallis, D. D., Putnam, E. A., Cretoiu, J. S., Carmical, S. G., Cao, S. N., Thomas, G., and Milewicz, D. M. (2003). Profibrillin-1 maturation by human dermal fibroblasts: Proteolytic processing and molecular chaperones. *J. Cell. Biochem.* **90,** 641–652.

Watson, R. E., Griffiths, C. E., Craven, N. M., Shuttleworth, C. A., and Kielty, C. M. (1999). Fibrillin-rich microfibrils are reduced in photoaged skin. Distribution at the dermal-epidermal junction. *J. Invest. Dermatol.* **112,** 782–788.

Werner, J. M., Knott, V., Handford, P. A., Campbell, I. D., and Downing, A. K. (2000). Backbone dynamics of a cbEGF domain pair in the presence of calcium. *J. Mol. Biol.* **296,** 1065–1078.

Wess, T. J., Purslow, P., and Kielty, C. M. (1997). Fibrillin-rich microfibrils: An X-ray diffraction study and elastomeric properties. *FEBS Lett.* **413,** 424–428.

Wess, T. J., Purslow, P. P., Sherratt, M. J., Ashworth, J. L., Shuttleworth, C. A., and Kielty, C. M. (1998a). Calcium determines the supramolecular organization of fibrillin-rich microfibrils. *J. Cell Biol.* **141,** 829–837.

Wess, T. J., Purslow, P. P., and Kielty, C. M. (1998b). X-ray diffraction studies of fibrillin-rich microfibrils: Effects of tissue extension on axial and lateral packing. *J. Struct. Biol.* **122,** 123–127.

Whiteman, P., and Handford, P. A. (2003). Defective secretion of recombinant fragments of fibrillin-1: Implications of protein misfolding for the pathogenesis of Marfan syndrome and related disorders. *Hum. Mol. Genet.* **12,** 727–737.

Yanagisawa, H., Davis, E. C., Starcher, B. C., Ouchi, T., Yanagisawa, M., Richardson, J. A., and Olsen, E. N. (2002). Fibulin-5 is an elastin-binding protein essential for elastic fibre development *in vivo. Nature* **415,** 168–171.

Yuan, X., Downing, A. K., Knott, V., and Handford, P. A. (1997). Solution structure of the transforming growth factor beta-binding protein-like module, a domain associated with matrix fibrils. *EMBO J.* **16,** 6659–6666.

Zhang, H., Apfelroth, S. D., Hu, W., Davis, E. C., Sanguineti, C., Bonadio, J., Mecham, R. P., and Ramirez, F. (1994). Structure and expression of fibrillin-2, a novel microfibrillar component preferentially located in elastic matrices. *J. Cell Biol.* **124,** 855–863.

Zhang, H., Hu, W., and Ramirez, F. (1995). Developmental expression of fibrillin genes suggests heterogeneity of extracellular microfibrils. *J. Cell Biol.* **129,** 1165–1176.

Zimmermann, D. R., Dours-Zimmermann, M., Schubert, M., and Bruckner-Tuderman, L. (1994). Versican is expressed in the proliferating zone in the epidermis and in association with the elastic network of the dermis. *J. Cell Biol.* **124,** 817–825.

ELASTIN

By SUZANNE M. MITHIEUX AND ANTHONY S. WEISS

School of Molecular and Microbial Biosciences, University of Sydney,
New South Wales 2006, Australia

I. Elastic Fiber 438
  A. Elastin 439
  B. Microfibrils 439
II. Elastogenesis 440
  A. Elastin Gene 440
  B. Regulation of Tropoelastin Expression 442
  C. Secretion 443
  D. Incorporation of Tropoelastin into the Elastic Fiber 444
  E. Coacervation 444
  F. Crosslinking 445
III. Properties of Elastin 446
  A. Physical Properties 446
  B. Structural Properties 447
  C. Biological Properties 448
IV. Mechanism of Elasticity 449
  A. Random Chain Model 449
  B. Liquid Drop Model 449
  C. Oiled Coil Model 449
  D. Fibrillar Model 450
V. Biomaterials 450
  References 453

Abstract

Elastin is a key extracellular matrix protein that is critical to the elasticity and resilience of many vertebrate tissues including large arteries, lung, ligament, tendon, skin, and elastic cartilage. Tropoelastin associates with multiple tropoelastin molecules during the major phase of elastogenesis through coacervation, where this process is directed by the precise patterning of mostly alternating hydrophobic and hydrophilic sequences that dictate intermolecular alignment. Massively crosslinked arrays of tropoelastin (typically in association with microfibrils) contribute to tissue structural integrity and biomechanics through persistent flexibility, allowing for repeated stretch and relaxation cycles that critically depend on hydrated environments. Elastin sequences interact with multiple proteins found in or colocalized with microfibrils, and bind to elastogenic cell surface receptors. Knowledge of the major stages in elastin assembly has facilitated the

ADVANCES IN PROTEIN CHEMISTRY, Vol. 70
DOI: 10.1016/S0065-3233(04)70013-3

Copyright 2005, Elsevier Inc.
All rights reserved.
0065-3233/05 $35.00

construction of *in vitro* models of elastogenesis, leading to the identification of precise molecular regions that are critical to elastin-based protein interactions.

I. Elastic Fiber

The extracellular matrix imparts structural integrity on the tissues and organs of the body. It also acts as a dynamic modulator of a variety of biological processes. An important component of the extracellular matrix is the elastic fiber. Elastic fibers confer the properties of elastic recoil and resilience on all vertebrate elastic tissues, with the exception of lower vertebrates such as the lamprey (Debelle and Tamburro, 1999). Such properties are critical to the long-term function of these tissues. Elastic fibers are found within arteries, lung, skin, vertebral *ligamenta flava*, vocal chords, and elastic cartilage (Sandberg *et al.*, 1981). They are composed of two morphologically and chemically distinct components–elastin and microfibrils. Elastin comprises approximately 90% of the elastic fiber and forms the internal core. It is interspersed with and surrounded by a sheath of unbranched microfibrils with an average diameter of 10–12 nm (Ramirez, 2000). The microfibrils are composed of a complex array of macromolecules.

The structure of elastic matrices differs among tissues. Their function is a consequence of their composition and organization or architecture. In arteries, the elastic fibers are organized into concentric rings of elastic lamellae around the arterial lumen (Fig. 1A). Each elastic lamella alternates with a ring of smooth muscle cells (Li *et al.*, 1998). In lung, elastic fibers form a delicate latticework throughout the organ, concentrating in areas of stress such as the opening of the alveoli and alveolar junctions (Fig. 1B). They provide the architectural foundation, as well as the stretch

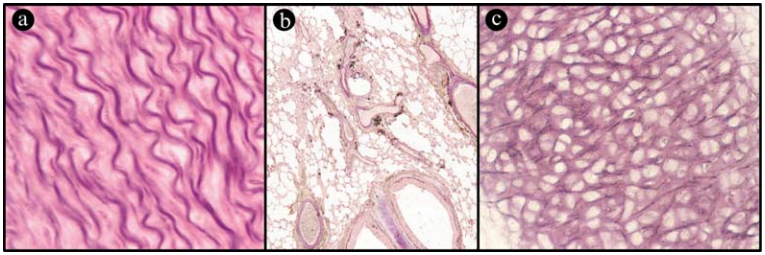


Fig. 1. The arrangement of elastin in (A) aorta, (B) lung, and (C) ear cartilage. Reprinted with permission from (A) Biodidac, (B) Dr. Alan Entwistle, Ludwig Institute for Cancer Research, and (C) Dr. Gwen Childs.

and recoil required for normal function (Starcher, 2000). In ligaments and tendons, the fibers are oriented parallel to the basic tissue organization. In elastic cartilage (Fig. 1C), the fibers are organized into a large three-dimensional honeycomb configuration (Dietz *et al.*, 1994).

A. *Elastin*

Elastin is an extremely durable, insoluble biopolymer that does not turn over appreciably in healthy tissue. It is estimated to have a half-life of about 70 years (Petersen *et al.*, 2002). Elastin is formed through the lysine-mediated crosslinking of its soluble precursor tropoelastin. Tropoelastin is an approximately 60–70 kDa protein, whose length depends on alternate splicing. Tropoelastin exists as a monomer in solution in two forms: an open globular molecule and a distended polypeptide (Toonkool *et al.*, 2001a). Crosslinking of tropoelastin is initiated through the action of the enzyme lysyl oxidase (Kagan and Sullivan, 1982) or other family members such as the LOX-like proteins 1–4 (Liu *et al.*, 2004; Noblesse *et al.*, 2004). It is the physical properties of elastin that are considered principally responsible for the function of elastic tissues (Keeley *et al.*, 2002). Elastin constitutes 30–57% of the aorta, 50% of elastic ligaments, 28–32% of major vascular vessels, 3–7% of lung, 4% of tendons, and 2–5% of the dry weight of skin (Vrhovski and Weiss, 1998).

B. *Microfibrils*

In developing elastic tissue, the microfibrils are the first components to appear in the extracellular matrix. They are then thought to act as a scaffold for deposition, orientation, and assembly of tropoelastin monomers. They are 10–12 nm in diameter, and lie adjacent to cells producing elastin and parallel to the long axis of the developing elastin fiber (Cleary, 1987).

Microfibrils are comprised of several macromolecules. Major components are the structural glycoproteins fibrillin-1 (Sakai *et al.*, 1986), fibrillin-2 (Zhang *et al.*, 1994), and microfibril-associated glycoprotein-1 (MAGP-1; Gibson *et al.*, 1986). The fibrillins are large (350 kDa), acidic, cysteine-rich glycoproteins that appear as extended, flexible molecules when viewed by electron microscopy. Their primary structure is dominated by calcium-binding epidermal growth factor-like repeats (Kielty *et al.*, 2002). MAGP-1 is an acidic, 31 kDa glycoprotein that is covalently linked to the microfibrils. Other proteins found associated with the microfibrils include vitronectin (Dahlback *et al.*, 1990), latent transforming growth factor β-binding

proteins (LTBPs; Kielty *et al.*, 2002), emilin (Bressan *et al.*, 1993; Mongiat *et al.*, 2000), microfibrillar associated proteins (MFAP), and members of the fibulin family (Nakamura *et al.*, 2002; Reinhardt *et al.*, 1996; Roark *et al.*, 1995; Yanagisawa *et al.*, 2002).

Proteoglycans such as versican (Isogai *et al.*, 2002), biglycan, and decorin (Reinboth *et al.*, 2002) interact with the microfibrils. They confer specific properties including hydration, impact absorption, molecular sieving, regulation of cellular activities, mediation of growth factor association, and release and transport within the extracellular matrix (Buczek-Thomas *et al.*, 2002). In addition, glycosaminoglycans have been shown to interact with tropoelastin through its lysine side chains (Wu *et al.*, 1999).

II. Elastogenesis

In vivo elastin fiber formation requires the coordination of a number of important processes. These include the control of intracellular transcription and translation of tropoelastin, intracellular processing of the protein, secretion of the protein into the extracellular space, delivery of tropoelastin monomers to sites of elastogenesis, alignment of the monomers with previously accreted tropoelastin through associating microfibrillar proteins, and finally, the conversion to the insoluble elastin polymer through the crosslinking action of lysyl oxidase (Fig. 2).

A. *Elastin Gene*

The human gene encoding tropoelastin is a single copy localized to 7q11.2 region (Fazio, 1991). The primary transcript is approximately 40 kb in length and contains small exons interspersed between large introns giving rise to an unusually low exon/intron ratio (Piontkivska *et al.*, 2004). This sequence codes for an mRNA of ~3.5 kb (Parks and Deak, 1990), which consists of a ~2.2 kb coding segment and a relatively large, 1.3 kb 3′ untranslated region (Rosenbloom *et al.*, 1991). The human tropoelastin gene contains 34 exons. Two further exons were lost in primate evolution through two steps, when exon 35 was probably deleted through inter-Alu recombination as Catarrhines diverged from Platyrrhines (New World monkeys). More recently, exon 34 was lost when *Homo sapiens* separated from the common ancestor shared with chimpanzees and gorillas. The additional exons persist in the bovine, porcine, feline, canine, and mouse genomes (Szabo *et al.*, 1999). Tropoelastin is distinguished by an exon periodicity where functionally distinct hydrophobic and crosslinking domains are encoded in separate alternating exons (Fig. 3). All the exons

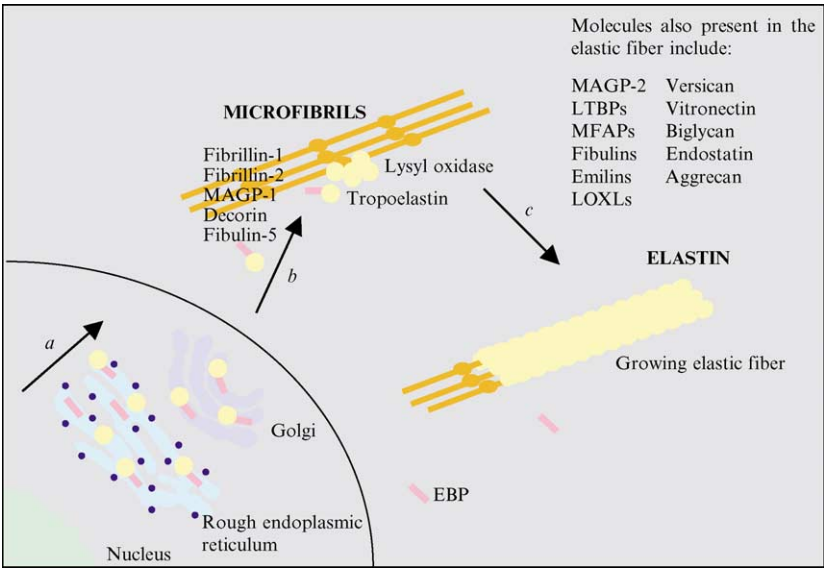


FIG. 2. Schematic representation of the classical model of elastogenesis. Tropoelastin is transcribed in mammals from a single gene and alternatively spliced in the nucleus. (A) Following translation and signal sequence cleavage, tropoelastin associates with EBP and FKBP65 in the rough endoplasmic reticulum. The tropoelastin-EBP complex then moves through the Golgi and is secreted to the cell surface. (B) Secreted tropoelastin is oxidized by a member of the lysyl oxidase family and tropoelastin associates with the microfibrils and with other tropoelastin molecules through coacervation to generate the nascent elastic fiber. (C) Continued secretion, oxidation, and depositing of tropoelastin occupy the bulk of elastin synthesis. The diagram is not drawn to scale. EBP, elastin binding protein; FKBP65, 65-kDa FK506 binding protein; MAGP, microfibril associated glycoprotein; LTBP, latent transforming growth factor β-binding protein; MFAP, microfibril associated protein; LOXL, lysyl oxidase like.

exist as multiples of three nucleotides and the exon-intron borders are always split in the same way (Rosenbloom *et al.*, 1993). The first nucleotide of a codon is found at the 3′ junction, while the second and third nucleotides are present at the 5′ border of exons. These exons are likely to have arisen through multiple intragenic duplication events (Kummerfeld *et al.*, 2003).

The primary transcript of tropoelastin undergoes extensive alternative splicing. As the splitting of codons at the exon-intron borders is consistent throughout the molecule, alternative splicing occurs in a cassette-like fashion with maintenance of the coding sequence (Bashir *et al.*, 1989).

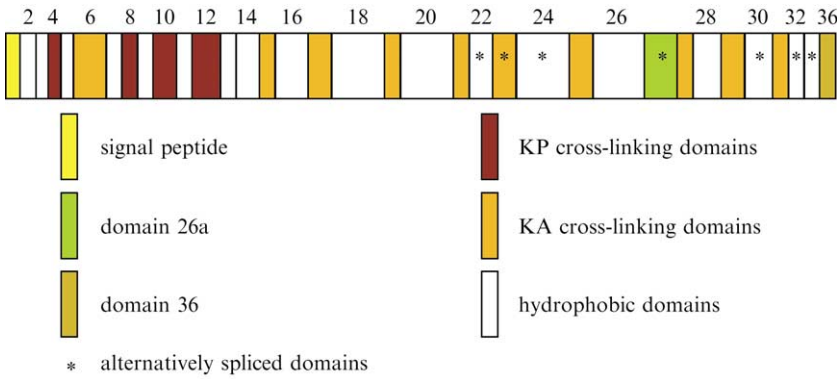


FIG. 3. Domain structure of human tropoelastin. Human tropoelastin consists of 34 domains and is dominated by alternating hydrophobic and crosslinking regions.

This results in the translation of multiple heterogeneous tropoelastin isoforms (Indik *et al.*, 1987). At least seven human exons are known to be alternatively spliced: 22, 23, 24, 26A, 30, 32, and 33 (Parks and Deak, 1990; Zhang M. C. *et al.*, 1999b). Alternative splicing of individual exons may be used to tailor the structural function of the protein in different tissues (Piontkivska *et al.*, 2004). It appears to be developmentally regulated and tissue-specific with age-related changes in isoform ratios observed in all species that have been investigated (Heim *et al.*, 1991; Parks and Deak, 1990; Starcher, 2000). The most frequently observed human tropoelastin isoform lacks exon 26A, which is reportedly only expressed in certain disease states (Debelle and Tamburro, 1999). Three human disorders have been linked to mutations or deletions of the tropoelastin gene: cutis laxa, supravalvular aortic stenosis, and Williams-Beuren syndrome.

B. Regulation of Tropoelastin Expression

Elastogenesis occurs primarily during late fetal and early neonatal periods. Elastin is synthesized and secreted from several cell types including smooth muscle cells, fibroblasts, endothelial cells, chondroblasts, and mesothelial cells (Uitto *et al.*, 1991) with tissue-specific induction of elastin expression during development (Swee *et al.*, 1995). After elastin has been deposited, its synthesis ceases and very little turnover of elastin is seen during adult life, unless the elastic fibers are subject to injury. In this case,

a program of neosynthesis of elastin can be rapidly activated (Davidson, 2002). Both pre- and posttranscriptional control mechanisms have been described for tropoelastin expression. The 5′-flanking region of the elastin gene is typical of many housekeeping genes; it contains GC-rich islands, no consensus TATA box, and multiple *cis*-elements that serve as potential sites for transcription-regulatory factors (Uitto *et al.*, 1991). Exogenous factors that alter elastin transcription include nuclear factor-1 family members (NF-1; Degterev and Foster, 1999), insulin-like growth factor-1 (IGF-1; Conn *et al.*, 1996; Wolfe *et al.*, 1993), basic fibroblast growth factor (bFGF; Carreras *et al.*, 2002; Rich *et al.*, 1999), tumor necrosis factor-α (Kahari *et al.*, 1992), interleukin 1 beta (Mauviel *et al.*, 1993), and interleukin 10 (Reitamo *et al.*, 1994).

Expression is extensively controlled at the posttranscriptional level (Parks, 1997), where mRNA decay is likely to be the predominant regulatory mechanism. Tropoelastin pre-mRNA levels remain elevated in mature adult lung, while steady-state mRNA levels are considerably reduced compared to those seen in neonatal lung (Swee *et al.*, 1995). Transforming growth factor (TGF-β1) is a strong stimulator of elastin expression and is believed to affect elastin mRNA stability (Liu and Davidson, 1988). It is thought to reduce the binding of a cytosolic, trans-acting protein whose interaction with mRNA destabilizes elastin transcripts (Davidson, 2002), although it is possible that microRNA is involved. A sequence coded by exon 30 of tropoelastin confers transcript stability and responsiveness to TGF-β1 (Zhang M. *et al.*, 1999a).

C. Secretion

After extensive splicing, the mature tropoelastin mRNA is exported from the nucleus. Translation occurs on the surface of the rough endoplasmic reticulum (rER) forming a 70-kDa polypeptide with an N-terminal signal sequence of 26 amino acids, which is cleaved as the protein enters the lumen of the rER (Grosso and Mecham, 1988). After release of the signal peptide, the protein travels through the lumen and is transported to the Golgi. Intracellularly, tropoelastin is likely to be chaperoned by a 67-kDa elastin-binding protein (EBP) that prevents intracellular self-aggregation and premature degradation (Hinek *et al.*, 1995). The EBP is an enzymatically inactive spliced variant of β-galactosidase and binds predominantly to hydrophobic domains on elastin (Hinek, 1995). FKBP65 is a peptidyl-prolyl *cis/trans* isomerase that also associates with tropoelastin in the secretory pathway (Patterson *et al.*, 2002).

D. *Incorporation of Tropoelastin into the Elastic Fiber*

After secretion, the EBP is likely to deliver tropoelastin to the microfibrillar scaffold site of fiber formation. The binding of microfibrillar galactosugars to the EBP results in the release of tropoelastin (Privitera *et al.*, 1998). Here it interacts with the microfibrils, which align the tropoelastin monomers, for subsequent formation of the elastic fiber. The formation of a specific transglutaminase crosslink between fibrillin-1 and tropoelastin may act to covalently stabilize the newly deposited tropoelastin on the microfibrils (Rock *et al.*, 2004), which is facilitated by calcium-dependent binding of MAGP-1 to multiple sites within tropoelastin (Clarke and Weiss, 2004). The C-terminal region of tropoelastin plays a pivotal role in the ordered deposition of the monomer into the growing elastin polymer (Brown-Augsburger *et al.*, 1996; Hsiao *et al.*, 1999; Kozel *et al.*, 2004). The EBP is recycled back to the intracellular endosomal compartments for reassociation with newly synthesized tropoelastin (Hinek *et al.*, 1995). The two major processes required for the incorporation of tropoelastin into the growing elastic fiber are coacervation and crosslinking.

E. *Coacervation*

In order for tropoelastin molecules to be crosslinked, they must first associate and align so as to facilitate the generation of crosslinks between closely spaced lysines. Coacervation is proposed to be the molecular mechanism through which aligning and concentrating can occur (Urry, 1978). Coacervation refers to an inverse temperature transition whereby tropoelastin molecules aggregate with increasing temperature. It is an entropically driven reversible process that involves interactions between the hydrophobic domains of tropoelastin. At low temperatures, tropoelastin is soluble in solution; on raising the temperature, the solution becomes cloudy as the tropoelastin molecules aggregate and become ordered by interactions between hydrophobic domains, such as the oligopeptide repetitive sequences GVGVP, GGVP, and GVGVAP (Vrhovski and Weiss, 1998). If left to settle, a viscoelastic phase forms containing highly concentrated tropoelastin. This phenomenon is explained in terms of the effects of water on the tropoelastin molecule. At low temperatures, water forms a clathrate-like structure around the hydrophobic regions of tropoelastin, keeping the protein unfolded. With increasing temperature, the ordered clathrate water is disrupted, rendering the hydrophobic domains free to fold and interact with other hydrophobic segments. Although order in the protein increases, the overall entropy of the system is increased

through the disruption of the water (Urry and Long, 1977; Urry *et al.*, 1969, 1974).

The process of coacervation is finely tuned to the physiological conditions of the extracellular matrix. Optimal coacervation of human tropoelastin occurs at 37 °C, 150 mM NaCl, and pH 7–8 (Vrhovski *et al.*, 1997). The arrangement of sequences in tropoelastin is critical to this process of coacervation, where association through hydrophobic domains depends on their contextual location in the molecule (Toonkool *et al.*, 2001b). Tropoelastin association rapidly proceeds through a monomer to *n*-mer transition, with little evidence of intermediate forms (Toonkool *et al.*, 2001a).

F. Crosslinking

Tropoelastin molecules are crosslinked in the extracellular space through the action of the copper-dependent amine oxidase, lysyl oxidase. Specific members of the lysyl oxidase-like family of enzymes are implicated in this process (Liu *et al.*, 2004; Noblesse *et al.*, 2004), although their direct roles are yet to be demonstrated enzymatically. Lysyl oxidase catalyzes the oxidative deamination of ϵ-amino groups on lysine residues (Kagan and Sullivan, 1982) within tropoelastin to form the α-aminoadipic-δ-semialdehyde, allysine (Kagan and Cai, 1995). The oxidation of lysine residues by lysyl oxidase is the only known posttranslational modification of tropoelastin. Allysine is the reactive precursor to a variety of inter- and intramolecular crosslinks found in elastin. These crosslinks are formed by nonenzymatic, spontaneous condensation of allysine with another allysine or unmodified lysyl residues. Crosslinking is essential for the structural integrity and function of elastin. Various crosslink types include the bifunctional crosslinks allysine-aldol and lysinonorleucine, the trifunctional crosslink merodesmosine, and the tetrafunctional crosslinks desmosine and isodesmosine (Umeda *et al.*, 2001).

Two types of crosslinking domains exist in tropoelastin: those rich in alanine (KA) and those rich in proline (KP). Within the KA domains, lysine residues are typically found in clusters of two or three amino acids, separated by two or three alanine residues. These regions are proposed to be α-helical with 3.6 residues per turn of helix, which has the effect of positioning two lysine sidechains on the same side of the helix, although there is no direct structural evidence (Brown-Augsburger *et al.*, 1995; Sandberg *et al.*, 1971), and facilitating the formation of desmosine crosslinks. Desmosine crosslinks are formed by the condensation of two allysine

residues on one tropoelastin and one allysine residue and one lysine residue on another molecule. In human tropoelastin, the KA domains are encoded by exons 6, 15, 17, 19, 21, 23, 25, 27, 29, and 31. The KP domains are encoded by exons 4, 8, 10, and 12. In these KP domains, the lysine pairs are separated by one or more proline residues and are flanked by prolines and bulky hydrophobic amino acids (Brown-Augsburger *et al.*, 1995). Desmosines and isodesmosines have not been found in association with KP domains. This is likely to be due to the steric constraints imposed by the presence of multiple proline residues that would not allow the formation of α-helical structure. Lysine residues are also found in domains 13, 14, and 36.

There is little information describing which lysine residues are involved in crosslink formation. This is largely due to the highly insoluble nature of elastin, making it difficult to analyze. However, one elastin crosslinking domain has been identified that joins three tropoelastin chains identified as 10, 19, and 25 (Brown-Augsburger *et al.*, 1995). The only two KA domains that contain three lysine residues are encoded by exons 19 and 25. They form a desmosine crosslink in an antiparallel arrangement of 19 and 25, while the remaining lysine on each chain forms a lysinonorleucine crosslink with two lysine residues present on the KP domain encoded by exon 10.

III. Properties of Elastin

A. *Physical Properties*

Purified elastin is pale yellow and has a characteristic blue fluorescence in ultraviolet light (Partridge, 1962). When dry, it is a hard, brittle glassy solid. On wetting, it becomes flexible and elastic. The water content of elastin is affected by temperature; a large increase is seen in the swollen volume of elastin with decreasing temperature (Lillie and Gosline, 2002). At 36 °C, purified bovine *ligamentum nuchae* (which is primarily composed of elastin) contains 0.46 g water/g protein; at 2 °C, this increases to 0.76 g water/g protein (Gosline, 1978; Gosline and French, 1979).

Young's Modulus provides a measure for the elasticity or stiffness of a material (i.e., the greater the Young's Modulus, the stiffer the material). The Modulus is determined from the slope of a stress/strain curve. The slope of this curve for elastin remains linear to an extension of ~70% (Gosline, 1976). The Young's Modulus of elastic fibers is 300–600 kPa; for collagen, it is 1×10^6 kPa. The maximum extension of elastic fibers ranges from 100–220% (Fung, 1993). Elastic fibers are able to undergo billions of

cycles of extension and recoil without mechanical failure (Keeley *et al.*, 2002).

B. *Structural Properties*

Numerous studies have been undertaken to elucidate the secondary structure of soluble elastin. These studies have been performed on elastin, elastin solubilized by oxalic acid (α-elastin) or potassium hydroxide (κ-elastin), synthetic polypeptide models of elastin, and tropoelastin. Techniques used include circular dichroism, FT-Raman, and electron microscopy. No consensus has been reached on the overall structure of elastin.

Circular dichroism (CD) studies on α-elastin (Tamburro *et al.*, 1977), κ-elastin, bovine, and human tropoelastin (Debelle *et al.*, 1995; Vrhovski *et al.*, 1997) have demonstrated a conformational transition to increased α-helical content with increasing temperature. The α-helical content predicted for tropoelastin is probably confined to the crosslinking domains, as the rest of the molecule is rich in helix breaking proline residues (Muiznieks *et al.*, 2003).

Urry *et al.* (1974) proposed that the hydrophobic regions of elastin containing the repeat peptides VPGVG, VPGG, and APGVGV form β-spirals on coacervation of the molecule. β-spirals contain recurring type II β-turns. Using CD analysis, Jensen *et al.* (2000) showed that domain 26 of human tropoelastin contained β-structure. Tamburro and colleagues have demonstrated that the hydrophobic domains of human tropoelastin may partly assume polyproline II (PPII) structures (Bochicchio *et al.*, 2004). Tropoelastin is likely to be a dynamic structure, as PPII does not contain obvious intramolecular hydrogen bonds and is likely to be in equilibrium with multiple conformations.

CD analysis of recombinant human tropoelastin shows that the molecule is composed of 3% α-helix, 41% β-sheet, 21% β-turn, and 33% other structure (Vrhovski *et al.*, 1997). FT-Raman studies on human elastin demonstrate derived secondary structures containing 8% α-helix, 36% β-strand, and 56% unordered conformation (Debelle *et al.*, 1998).

Macroscopically, elastin appears to be an amorphous mass. Ultrastructural electron microscopy studies reveal that elastin has a fibrillar substructure comprised of parallel-aligned $\sim$5 nm thick filaments that appear to have a twisted ropelike structure (Gotte *et al.*, 1974; Pasquali-Ronchetti *et al.*, 1998). A variety of techniques have been used to resolve these filaments, including negative staining electron microscopy of sonicated fragments of purified elastic fibers (Serafini-Fracassini *et al.*, 1976), freeze

fracture electron microscopy techniques on bovine *ligamentum nuchae* (Pasquali-Ronchetti *et al.*, 1979), ultrathin cryosections of natural fresh elastic fibers (Fornieri *et al.*, 1982), and aggregated tropoelastin molecules (Bressan *et al.*, 1986). Scanning force microscopy also confirmed the filamentary network organization of elastin. These results suggest that crosslinked elastin molecules exhibit globular domains preferentially linked in one direction to form filaments and laterally joined at regular intervals (Pasquali-Ronchetti *et al.*, 1998).

C. *Biological Properties*

The key biological function of elastin is to impart elasticity to organs and tissues. However, studies have demonstrated that elastin and elastin peptides have diverse biological properties (Faury *et al.*, 1998). Evidence for both direct and indirect elastin-mediated cell signaling has been demonstrated. Elastin regulates arterial and lung terminal airway branching morphogenesis (Li *et al.*, 1998; Wendel *et al.*, 2000). Elastin receptors include the elastin-laminin receptor (ELR; Hinek *et al.*, 1988), which relies on an EBP subunit to facilitate association, and integrin $\alpha v\beta 3$ (Rodgers and Weiss, 2004). The ELR mediates the regulation of skin fibroblast proliferation (Groult *et al.*, 1991), chemotaxis for monocytes and fibroblasts (Indik *et al.*, 1990; Senior *et al.*, 1980, 1982), age-dependent vasodilation in aortic rings (Faury *et al.*, 1995), endothelium-dependent vasorelaxation (Faury *et al.*, 1998), inhibition of the migratory response of smooth muscle cells (SMC) to chemoattractants (Ooyama *et al.*, 1987), regulation of SMC proliferation (Ito *et al.*, 1998), and the increase of Ca^{2+} levels in leukocytes and endothelial cells (Faury *et al.*, 1998; Varga *et al.*, 1989). Elastin-derived protein coating of a poly(ethylene terephthalate), commonly used in biomaterials, promotes the growth, proliferation, and phenotype maintenance of endothelial cells (Dutoya *et al.*, 2000).

The repetitive mechanical deformation in arterial tissues affects phenotype and proliferation of smooth muscle cells (Birukov *et al.*, 1995). The physiological force of wall stress is a key determinant of arterial development (Li *et al.*, 1998) and cyclic stretching stimulates the synthesis of extracellular matrix components by arterial smooth muscle cells *in vitro* (Leung *et al.*, 1976). Furthermore, the development of mechanically functional small-caliber autologous arteries by seeding biodegradable polyglycolic acid (PGA) tube scaffolds with SMCs on the exterior and endothelial cells on the interior have been limited due to a lack of elastin deposition (Niklason *et al.*, 1999). These data collectively point to the importance of an exogenous supply of elastin-like materials in biomaterial constructs (van Hest and Tirrell, 2001).

IV. Mechanism of Elasticity

The fundamental driving force behind the remarkable elastic properties of the elastin polymer is believed to be entropic, where stretching decreases the entropy of the system and elastic recoil is driven by a spontaneous return to maximum entropy. The precise molecular basis for elasticity has not been fully elucidated and a number of models exist. Two main categories of structure-function models have been proposed: those in which elastin is considered to be isotropic and devoid of structure, and those which consider elastin to be anisotropic with regions of order (Vrhovski and Weiss, 1998).

A. *Random Chain Model*

This model is based on classic rubber theory, suggesting that elastin is made up of a network of random chains that are kinetically free and exist in a high entropic state. Stretching orders the chains and limits their conformational freedom, thus decreasing the overall entropy of the system (Hoeve and Flory, 1974). This provides the restoring force to the relaxed state.

B. *Liquid Drop Model*

In this model, the hydration of hydrophobic side chains of tropoelastin, which become exposed to solvent on extension, is implicated in the stretch-induced decrease in entropy (Urry *et al.*, 2002; Weis-Fogh and Andersen, 1970). Weis-Fogh and Andersen (1970) proposed that globular, spherical monomers (oil droplets) of tropoelastin are crosslinked to form a three-dimensional aggregate. Deformation of the matrix exposes the interior nonpolar groups to water. Gosline (1980) suggested a large hydrophobic contribution to the stored elastic energy; however, a random network arrangement of tropoelastin is implied in this model.

C. *Oiled Coil Model*

This model is similar to the liquid drop model but regards the tropoelastin monomer as fibrillar. It features a broad coil made up of the repeating tetrapeptide units GVPG in which glycine residues occupy the exterior portions exposed to solvent, while proline, valine, and other hydrophobic residues are buried. Each monomer is fibrillar, made up of alternating sections of crosslink regions and ''oiled coils'' (Gray *et al.*, 1973). The

crosslink regions are rigid, while the oiled coil hydrophobic regions are flexible. On extension, as for the liquid drop model, the hydrophobic interior is exposed to water providing an entropic change in the system.

D. *Fibrillar Model*

The fibrillar model (Urry *et al.*, 1974) proposes that elasticity arises from the properties of regular β-spiral structures comprising repetitive β-turns. These stable β-turns are formed by regularly repeating sequences (e.g., VPGVG, VPGG, and APGVGV) that are found in tropoelastin. The β-turns act as spacers between the turns of the β-spiral, suspending chain segments in a conformationally free state. On stretching, the liberational entropy in these peptide segments is reduced, which provides the restoring force required for elasticity.

Tamburro *et al.* (1991) also proposed a β-turn based model, where the β-turns are situated in the GXGGX repeating sequences. These glycine turns are considered to be more labile than proline counterparts and able to interconvert, giving rise to dynamic β-turns sliding along the protein chain (Debelle and Tamburro, 1999). In this model, the polypeptide chain is freely fluctuating, and the system is essentially unstructured. The entropic restoring force is considered to be similar to that proposed by the random chain model.

V. Biomaterials

In healthy individuals, the mature elastin molecule is a stable, insoluble protein. Degradation of elastin is extremely slow due to the extensive crosslinking of tropoelastin within the elastic fiber. However, with aging, injury, or the onset of a variety of acquired diseases, the degradation and excessive or aberrant remodeling of elastic fibers becomes apparent in arteries, lung, skin, and ligament (Osakabe *et al.*, 2001). The correct assembly of the elastin polymer is critical for proper elastic function, and therefore the physiological properties of these tissues become compromised. In aortic valves, damage to elastin leads to a passive elongation of the tissue, a reduction in extensibility, and an increase in stiffness (Lee T. C. *et al.*, 2001b). Examples of common disorders that involve abnormal elastic fiber assembly and/or degradation include aortic aneurysms, atherosclerosis, chronic obstructive pulmonary disease, emphysema, and hypertension (Urban and Boyd, 2000). In addition, a range of genetic disorders can affect the elastic fiber system, leading to skeletal and skin abnormalities and vascular defects (Milewicz *et al.*, 2000).

Elastic fiber degradation due to age or disease is caused by the action of a group of proteolytic enzymes, known as elastases (Faury, 2001). Elastolytic enzymes belong to four classes of proteases: serine, aspartic, cysteine, and metalloproteinases (Jacob *et al.*, 2001). Accompanying the degradation of elastin is the presence of a significant amount of degradation products in the tissues and circulation system. These elastin peptides can induce diverse biological responses, including extracellular matrix protein synthesis and deposition, cell attachment, migration, and proliferation.

The need to repair or replace damaged tissues and organs is being increasingly met through the development of innovative biomaterials. Ideal implants should mimic the native cellular environment and possess appropriate mechanical properties, such as strength and elasticity. Consequently, extracellular matrix engineering has become a major area of interest in the field of tissue repair (Stock *et al.*, 2001). Almost all somatic cells are in contact with the complex network of macromolecules that make up the extracellular matrix. This matrix is particularly plentiful in connective tissues where cells are only sparsely distributed (Sittinger *et al.*, 1996). Elastin-mediated elasticity is crucial for the normal function of skin, lung, bladder, and arteries.

Multiple approaches to tissue replacement and guided tissue regeneration are currently being used. These include using autografts, allografts, synthetically or naturally derived biomaterials, and cell:matrix constructs where cells are placed on or within the matrix (Fuchs *et al.*, 2001). Regardless of the approach, the necessary criteria for a suitable tissue replacement system are that it is biocompatible, bioactive, and biofunctional.

The requirement for elastic behavior in many replacement tissues, and the diverse biological properties elicited by elastin, have led to increased interest in the development of artificial biodegradable elastomers such as hydrogels (Hoffman, 2001), elastin-peptide based or elastin-mimetic polymers (Cappello *et al.*, 1998; Dutoya *et al.*, 2000; Huang *et al.*, 2000; Keeley *et al.*, 2002; Martino and Tamburro, 2001; Martino *et al.*, 2002; McMillan and Conticello, 2000; Panitch *et al.*, 1999; Rovira *et al.*, 1996; Urry *et al.*, 1998), and devitalized exogenous elastin matrices from bovine, porcine, and canine sources (Goissis *et al.*, 2000; Kajitani *et al.*, 2000, 2001; Vardaxis *et al.*, 1996). Proposed uses for elastin-based materials include prostheses for vascular walls, calcifiable matrices that could induce ossification in areas where bone formation is desired (Urry, 1978; Vardaxis *et al.*, 1996), arterial graft coating to exploit the antiproliferative effect on smooth muscle cells (Ito *et al.*, 1998), matrices for *in situ* formation of cartilaginous tissue (Betre *et al.*, 2002) and targeted drug delivery (Cappello *et al.*, 1998; Chilkoti *et al.*, 2002).

Elastomers are commonly used in applications that require compliance with soft or cardiovascular tissues (Peppas and Langer, 1994). Hydrogels are polymer networks that may absorb from 10–20% up to thousands of times their dry weight in water (Hoffman, 2001). Elastomeric polypeptides crosslinked into matrices absorb 400–600% of their dry weight in water (Lee J. *et al.*, 2001a). Hydrogels are finding increased use in medicine and biomedical engineering because of their capacity for rapid solute transfer and physical resemblance to living tissues (Martin *et al.*, 1998). Investigations into the use of stimuli-sensitive hydrogels that sense environmental change, causing an induction of structural change to the material, are mounting due to their potential applications in the development of biomaterials and drug delivery systems (Miyata *et al.*, 2002).

The first elastin-based polymers contained the polypeptide sequence thought to be involved in the elastic properties of elastin: poly(GVGVP). These polymers were shown to exist in hydrogel, elastic, and plastic states (Guda *et al.*, 1995). More recently, polymers that contain a variety of subdomains of elastin based on this sequence have been shown to exhibit self-assembling properties and lower critical solution temperature behavior. They are referred to as *stimuli responsive polymers* (Nath and Chilkoti, 2002). While these polymers display worthwhile mechanical properties, they lack many other elastin sequences of biological relevance. Consequently, they require the incorporation of epitopes that are directly bound by cellular adhesion receptors, such as RGD (Kobatake *et al.*, 2000; Nicol *et al.*, 1992) or REDV (Panitch *et al.*, 1999), which stimulate endothelial cell binding. These approaches are being increasingly investigated as they are expected to encourage cellular infiltration and allow for remodeling and incorporation of elastin-based polymers as living tissue (Kajitani *et al.*, 2001). Remarkably, through its natural C-terminus, tropoelastin is now appreciated to bind directly to integrin $\alpha v\beta 3$, which raises the opportunity of using elastin sequences instead of hybrid constructs to facilitate cell interactions (Rodgers and Weiss, 2004).

While the use of exogenous, devitalized elastin purified from animal sources may be limited due to problems associated with prion-mediated spongiform encephalopathies such as Jakob-Creutzfeldt disease, investigations into this source of elastin are still being considered. Partially devitalized collagen and elastin matrices purified from blood vessels obtained from beagle dogs are being examined for their potential use in small-diameter vascular grafts (Goissis *et al.*, 2000). An acellular elastin patch made from porcine aorta has been used to successfully repair esophageal and duodenal injuries in pigs (Kajitani *et al.*, 2001, 2000). Elastin-solubilized peptides from bovine *ligamentum nuchae* have been

chemically bonded to Dacron (polyethylene terephthalate) to elicit biological activity and improve compatibility (Bonzon *et al.*, 1995).

An elastic polymer matrix was made using as few as three hydrophobic domains flanking two crosslinking domains (Bellingham *et al.*, 2003). Recently, a synthetic elastin (Mithieux *et al.*, 2004) was made *in vitro* through the chemical crosslinking of recombinant human tropoelastin (Martin *et al.*, 1995). The synthetic elastin displays physical, structural, and biological properties similar to those of naturally occurring human elastin. These findings define the *in vitro* system as a useful model for elastogenesis. Synthetic elastin has potential as a novel biomaterial that is easily manufactured, can be molded into a variety of shaped tissue substrates, and has a range of properties that are required for elastic, compliant, cell-interacting, and medically relevant applications. Knowledge of many of the molecules that interact with tropoelastin and elastin, combined with the development of *in vitro* model elastogenic systems, demonstrate the versatility of elastin and facilitate the development of a new generation of biomaterials based on synthetic human elastin that can be considered as scaffolds for the augmentation and repair of human tissue.

Acknowledgments

ASW is a recipient of grants from the Australian Research Council and the University of Sydney Vice Chancellor's Development Fund.

References

Bashir, M. M., Indik, Z., Yeh, H., Ornstein-Goldstein, N., Rosenbloom, J. C., Abrams, W., Fazio, M., Uitto, J., and Rosenbloom, J. (1989). Characterization of the complete human elastin gene. Delineation of unusual features in the 5′-flanking region. *J. Biol. Chem.* **264,** 8887–8891.

Bellingham, C. M., Lillie, M. A., Gosline, J. M., Wright, G. M., Starcher, B. C., Bailey, A. J., Woodhouse, K. A., and Keeley, F. W. (2003). Recombinant human elastin polypeptides self-assemble into biomaterials with elastin-like properties. *Biopolymers* **70,** 445–455.

Betre, H., Setton, L. A., Meyer, D. E., and Chilkoti, A. (2002). Characterization of a genetically engineered elastin-like polypeptide for cartilaginous tissue repair. *Biomacromolecules* **3,** 910–916.

Birukov, K. G., Shirinsky, V. P., Stepanova, O. V., Tkachuk, V. A., Hahn, A. W. A., Resink, T. J., and Smirnov, V. N. (1995). Stretch affects phenotype and proliferation of vascular smooth muscle cells. *Mol. Cell. Biochem.* **144,** 131–139.

Bochicchio, B., Aït-Ali, A., Tamburro, A. M., and Alix, A. J. P. (2004). Spectroscopic evidence revealing polyproline II structure in hydrophobic, putatively elastomeric sequences encoded by specific exons of human tropoelastin. *Biopolymers* **73,** 484–493.

Bonzon, N., Lefebvre, F., Ferre, N., Daculsi, G., and Rabaud, M. (1995). New bioactivation mode for vascular prostheses made of Dacron polyester. *Biomaterials* **16,** 747–751.

Bressan, G. M., Pasquali-Ronchetti, I., Fornieri, C., Mattioli, F., Castellani, I., and Volpin, D. (1986). Relevance of aggregation properties of tropoelastin to the assembly and structure of elastic fibers. *J. Ultrastruct. Molec. Struct. Res.* **94,** 209–216.

Bressan, G. M., Daga-Gordini, D., Colombatti, A., Castellani, I., Marigo, V., and Volpin, D. (1993). Emilin, a component of elastic fibers preferentially located at the elastin-microfibrils interface. *J. Cell Biol.* **121,** 201–212.

Brown-Augsburger, P., Tisdale, C., Broekelmann, T., Sloan, C., and Mecham, R. P. (1995). Identification of an elastin cross-linking domain that joins three peptide chains. Possible role in nucleated assembly. *J. Biol. Chem.* **270,** 17778–17783.

Brown-Augsburger, P. B., Broekelmann, T., Rosenbloom, J., and Mecham, R. P. (1996). Functional domains on elastin and MAGP involved in elastic fiber assembly. *Biochem. J.* **318,** 149–155.

Buczek-Thomas, J. A., Chu, C. L., Rich, C. B., Stone, P. J., Foster, J. A., and Nugent, M. A. (2002). Heparan sulfate depletion within pulmonary fibroblasts: Implications for elastogenesis and repair. *J. Cell. Physiol.* **192,** 294–303.

Cappello, J., Crissman, J. W., Crissman, M., Ferrari, F. A., Textor, G., Wallis, O., Whitledge, J. R., Zhou, X., Burman, D., Aukerman, L., and Stedronsky, E. R. (1998). *In-situ* self-assembling protein polymer gel systems for administration, delivery, and release of drugs. *J. Control. Release* **53,** 105–117.

Carreras, I., Rich, C. B., Panchenko, M. P., and Foster, J. A. (2002). Basic fibroblast growth factor decreases elastin gene transcription in aortic smooth muscle cells. *J. Cell. Biochem.* **85,** 592–600.

Chilkoti, A., Dreher, M. R., Meyer, D. E., and Raucher, D. (2002). Targeted drug delivery by thermally responsive polymers. *Adv. Drug Deliver. Rev.* **54,** 613–630.

Clarke, A. W., and Weiss, A. S. (2004). Microfibril-associated glycoprotein-1 binding to tropoelastin: Multiple binding sites and the role of divalent cations. *Eur. J. Biochem.* **271,** 3085–3090.

Cleary, E. G. (1987). The microfibrillar component of the elastic fibers. *In* ''Connective Tissue Disease. Molecular Pathology of the Extracellular Matrix'' (J. Uitto and A. J. Perejda, Eds.), pp. 55–81. Dekker, New York.

Conn, K. J., Rich, C. B., Jensen, D. E., Fontanilla, M. R., Bashir, M. M., Rosenbloom, J., and Foster, J. A. (1996). Insulin-like growth factor-I regulates transcription of the elastin gene through a putative retinoblastoma control element. A role for Sp3 acting as a repressor of elastin gene transcription. *J. Biol. Chem.* **271,** 28853–28860.

Dahlback, K., Ljungquist, A., Lofberg, H., Dahlback, B., Engvall, E., and Sakai, L. Y. (1990). Fibrillin immunoreactive fibers constitute a unique network in the human dermis: Immunohistochemical comparison of the distributions of fibrillin, vitronectin, amyloid P component, and orcein stainable structures in normal skin and elastosis. *J. Invest. Dermatol.* **94,** 284–291.

Davidson, J. M. (2002). Smad about elastin regulation. *Am. J. Resp. Cell Mol.* **26,** 164–166.

Debelle, L., Alix, A. J., Jacob, M. P., Huvenne, J. P., Berjot, M., Sombret, B., and Legrand, P. (1995). Bovine elastin and kappa-elastin secondary structure determination by optical spectroscopies. *J. Biol. Chem.* **270,** 26099–26103.

Debelle, L., Alix, A. J., Wei, S. M., Jacob, M. P., Huvenne, J. P., Berjot, M., and Legrand, P. (1998). The secondary structure and architecture of human elastin. *Eur. J. Biochem.* **258,** 533–539.

Debelle, L., and Tamburro, A. M. (1999). Elastin: Molecular description and function. *Int. J. Biochem. Cell Biol.* **31,** 261–272.

Degterev, A., and Foster, J. A. (1999). The role of NF-1 factors in regulation of elastin gene transcription. *Matrix Biol.* **18,** 295–307.

Dietz, H. C., Ramirez, F., and Sakai, L. Y. (1994). Marfan's syndrome and other microfibrillar diseases. *Adv. Hum. Genet.* **22,** 153–186.

Dutoya, S., Verna, A., Lefebvre, F., and Rabaud, M. (2000). Elastin-derived protein coating onto poly(ethylene terephthalate). Technical, microstructural and biological studies. *Biomaterials* **21,** 1521–1529.

Faury, G., Ristori, M. T., Verdetti, J., Jacob, M. P., and Robert, L. (1995). Effect of elastin peptides on vascular tone. *J. Vascul. Res.* **32,** 112–119.

Faury, G., Garnier, S., Weiss, A. S., Wallach, J., Fulop, T., Jr, Jacob, M. P., Mecham, R. P., Robert, L., and Verdetti, J. (1998). Action of tropoelastin and synthetic elastin sequences on vascular tone and on free Ca^{2+} level in human vascular endothelial cells. *Circ. Res.* **82,** 328–336.

Faury, G. (2001). Function-structure relationship of elastic arteries in evolution: From microfibrils to elastin and elastic fibers. *Pathol. Biol.* **49,** 310–325.

Fazio, M. J., Mattei, M.-G., Passage, E., Chu, M.-L., Black, D., Solomon, E., Davidson, J. M., and Uitto, J. (1991). Human elastin gene: New evidence for localization to the long arm of chromosome 7. *Am. J. Hum. Genet.* **48,** 696–703.

Fornieri, C., Ronchetti, I. P., Edman, A. C., and Sjostrom, M. (1982). Contribution of cryotechniques to the study of elastin ultrastructure. *J. Microsc.* **126,** 87–93.

Fuchs, J. R., Nasseri, B. A., and Vacanti, J. P. (2001). Tissue engineering: A 21st century solution to surgical reconstruction. *Ann. Thor. Surg.* **72,** 577–591.

Fung, Y. C. (1993). Biomechanics. Mechanical properties of living tissues. Springer-Verlag, New York.

Gibson, M. A., Hughes, J. L., Fanning, J. C., and Cleary, E. G. (1986). The major antigen of elastin-associated microfibrils is a 31-kDa glycoprotein. *J. Biol. Chem.* **261,** 11429–11436.

Goissis, G., Suzigan, S., Parreira, D. R., Maniglia, J. V., Braile, D. M., and Raymundo, S. (2000). Preparation and characterization of collagen-elastin matrices from blood vessels intended as small diameter vascular grafts. *Artif. Organs* **24,** 217–223.

Gosline, J. M. (1976). The physical properties of elastic tissue. *Int. Rev. Connect. Tissue Res.* **7,** 211–249.

Gosline, J. M. (1978). The temperature-dependent swelling of elastin. *Biopolymers* **17,** 697–707.

Gosline, J. M., and French, C. J. (1979). Dynamic mechanical properties of elastin. *Biopolymers* **18,** 2091–2103.

Gosline, J. M. (1980). The elastic properties of rubber-like proteins and highly extensible tissues. *Sym. Soc. Exp. Biol.* **34,** 332–357.

Gotte, L., Giro, M. G., Volpin, D., and Horne, R. W. (1974). The ultrastructural organization of elastin. *J. Ultrastruct. Res.* **46,** 23–33.

Gray, W. R., Sandberg, L. B., and Foster, J. A. (1973). Molecular model for elastin structure and function. *Nature* **246,** 461–466.

Grosso, L. E., and Mecham, R. P. (1988). *In vitro* processing of tropoelastin: Investigation of a possible transport function associated with the carboxy-terminal domain. *Biochem. Biophys. Res. Comm.* **153,** 545–551.

Groult, V., Hornebeck, W., Ferrari, P., Tixier, J. M., Robert, L., and Jacob, M. P. (1991). Mechanisms of interaction between human skin fibroblasts and elastin: Differences between elastin fibers and derived peptides. *Cell Biochem. Funct.* **9,** 171–182.

Guda, C., Zhang, X., McPherson, D. T., Xu, J., Cherry, J. H., Urry, D. W., and Daniell, H. (1995). Hyper expression of an environmentally friendly synthetic polymer gene. *Biotech. Lett.* **17,** 745–750.

Heim, R. A., Pierce, R. A., Deak, S. B., Riley, D. J., Boyd, C. D., and Stolle, C. A. (1991). Alternative splicing of rat tropoelastin mRNA is tissue-specific and developmentally regulated. *Matrix* **11,** 359–366.

Hinek, A., Wrenn, D. S., Mecham, R. P., and Barondes, S. H. (1988). The elastin receptor: A galactoside-binding protein. *Science* **239,** 1539–1541.

Hinek, A. (1995). The 67 kDa spliced variant of beta-galactosidase serves as a reusable protective chaperone for tropoelastin. *Ciba Found. Sym.* **192,** 185–191.

Hinek, A., Keeley, F. W., and Callahan, J. (1995). Recycling of the 67-kDa elastin binding protein in arterial myocytes is imperative for secretion of tropoelastin. *Exp. Cell Res.* **220,** 312–324.

Hoeve, C. A., and Flory, P. J. (1974). The elastic properties of elastin. *Biopolymers* **13,** 677–686.

Hoffman, A. S. (2001). Hydrogels for biomedical applications. *Ann. N. Y. Acad. Sci.* **944,** 62–73.

Hsiao, H., Stone, P. J., Toselli, P., Rosenbloom, J., Franzblau, C., and Schreiber, B. M. (1999). The role of the carboxy terminus of tropoelastin in its assembly into the elastic fiber. *Connect. Tissue Res.* **40,** 83–95.

Huang, L., McMillan, R. P. A., Pourdeyhimi, B., Conticello, V. P., and Chaikof, E. L. (2000). Generation of synthetic elastin-mimetic small diameter fibers and fiber networks. *Macromolecules* **33,** 2989–2997.

Isogai, Z., Aspberg, A., Keene, D. R., Ono, R. N., Reinhardt, D. P., and Sakai, L. Y. (2002). Versican interacts with fibrillin-1 and links extracellular microfibrils to other connective tissue networks. *J. Biol. Chem.* **277,** 4565–4572.

Indik, Z., Yeh, H., Ornstein-Goldstein, N., Sheppard, P., Anderson, N., Rosenbloom, J. C., Peltonen, L., and Rosenbloom, J. (1987). Alternative splicing of human elastin mRNA indicated by sequence analysis of cloned genomic and complementary DNA. *Proc. Natl. Acad. Sci. USA* **84,** 5680–5684.

Indik, Z., Abrams, W. R., Kucich, U., Gibson, C. W., Mecham, R. P., and Rosenbloom, J. (1990). Production of recombinant human tropoelastin: Characterization and demonstration of immunologic and chemotactic activity. *Arch. Biochem. Biophys.* **280,** 80–86.

Ito, S., Ishimaru, S., and Wilson, S. E. (1998). Effect of coacervated alpha-elastin on proliferation of vascular smooth muscle and endothelial cells. *Angiology* **49,** 289–297.

Jacob, M. P., Badier-Commander, C., Fontaine, V., Benazzoug, Y., Feldman, L., and Michel, J. B. (2001). Extracellular matrix remodeling in the vascular wall. *Pathol. Biol.* **49,** 326–332.

Jensen, S. A., Vrhovski, B., and Weiss, A. S. (2000). Domain 26 of tropoelastin plays a dominant role in association by coacervation. *J. Biol. Chem.* **275,** 28449–28454.

Kagan, H. M., and Sullivan, K. A. (1982). Lysyl oxidase: Preparation and role in elastin biosynthesis. *Methods Enzymol.* **82,** 637–650.

Kagan, H. M., and Cai, P. (1995). Isolation of active site peptides of lysyl oxidase. *Methods Enzymol.* **258,** 122–132.

Kahari, V. M., Chen, Y. Q., Bashir, M. M., Rosenbloom, J., and Uitto, J. (1992). Tumor necrosis factor-alpha down-regulates human elastin gene expression. Evidence for the role of AP-1 in the suppression of promoter activity. *J. Biol. Chem.* **267,** 26134–26141.

Kajitani, M., Wadia, Y., Xie, H., Hinds, M. T., Shalaby, S. W., Swartz, K. R., and Gregory, K. W. (2000). Use of a new elastin patch and glue for repair of a major duodenal injury. *Am. Soc. Art. Int. Organs J.* **46,** 409–414.

Kajitani, M., Wadia, Y., Hinds, M. T., Teach, J., Swartz, K. R., and Gregory, K. W. (2001). Successful repair of esophageal injury using an elastin based biomaterial patch. *J. Am. Soc. Art. Int. Organs* **47,** 342–345.

Keeley, F. W., Bellingham, C. M., and Woodhouse, K. A. (2002). Elastin as a self-organising biomaterial: Use of recombinantly expressed human elastin polypeptides as a model for investigations of structure and self-assembly of elastin. *Philosoph. Trans. Royal Soc. Lond. B.* **357,** 185–189.

Kielty, C. M., Sherratt, M. J., and Shuttleworth, C. A. (2002). Elastic fibers. *J. Cell Sci.* **115,** 2817–2828.

Kobatake, E., Onoda, K., Yanagida, Y., and Aizawa, M. (2000). Design and gene engineering synthesis of an extremely thermostable protein with biological activity. *Biomacromolecules* **1,** 382–386.

Kozel, B.A, Ciliberto, C. H., and Mecham, R. P. (2004). Deposition of tropoelastin into the extracellular matrix requires a competent elastic fiber scaffold but not live cells. *Matrix Biol.* **23,** 23–34.

Kummerfeld, S. K., Weiss, A. S., Fekete, A., and Jermiin, L. S. (2003). AMID: Autonomous modeler of intragenic duplication. *Appl. Bioinform.* **2,** 169–176.

Lee, J., Macosko, C. W., and Urry, D. W. (2001a). Mechanical properties of cross-linked synthetic elastomeric polypentapeptides. *Macromolecules* **34,** 5968–5974.

Lee, T. C., Midura, R. J., Hascall, V. C., and Vesely, I. (2001b). The effect of elastin damage on the mechanics of the aortic valve. *J. Biomech.* **34,** 203–210.

Leung, D. Y. M., Glagov, S., and Mathews, M. B. (1976). Cyclic stretching stimulates synthesis of matrix components by arterial smooth muscle cells *in vitro. Science* **191,** 475–477.

Li, D. Y., Brooke, B., Davis, E. C., Mecham, R. P., Sorensen, L. K., Boak, B. B., Eichwald, E., and Keating, M. T. (1998). Elastin is an essential determinant of arterial morphogenesis. *Nature* **393,** 276–280.

Lillie, M. A., and Gosline, J. M. (2002). Unusual swelling of elastin. *Biopolymers* **64,** 115–126.

Liu, J. M., and Davidson, J. M. (1988). The elastogenic effect of recombinant transforming growth factor-beta on porcine aortic smooth muscle cells. *Biochem. Biophys. Res. Comm.* **154,** 895–901.

Liu, X., Zhao, Y., Gao, J., Pawlyk, B., Starcher, B., Spencer, J. A., Yanagisawa, H., Zuo, J., and Li, T. (2004). Elastic fiber homeostasis requires lysyl oxidase-like 1 protein. *Nat. Genet.* **36,** 178–182.

Martin, B. D., Linhardt, R. J., and Dordick, J. S. (1998). Highly swelling hydrogels from ordered galactose-based polyacrylates. *Biomaterials* **19,** 69–76.

Martin, S. L., Vrhovski, B., and Weiss, A. S. (1995). Total synthesis and expression in Escherichia coli of a gene encoding human tropoelastin. *Gene.* **154,** 159–166.

Martino, M., and Tamburro, A. M. (2001). Chemical synthesis of cross-linked poly(KGGVG), an elastin-like biopolymer. *Biopolymers* **59,** 29–37.

Martino, M., Perri, T., and Tamburro, A. M. (2002). Elastin-based biopolymers: Chemical synthesis and structural characterization of linear and cross-linked poly (OrnGlyGlyOrnGly). *Biomacromolecules* **3,** 297–304.

Mauviel, A., Chen, Y. Q., Kahari, V. M., Ledo, I., Wu, M., Rudnicka, L., and Uitto, J. (1993). Human recombinant interleukin-1 beta up-regulates elastin gene expression

in dermal fibroblasts. Evidence for transcriptional regulation *in vitro* and *in vivo*. *J. Biol. Chem.* **268,** 6520–6524.

McMillan, R. A., and Conticello, V. P. (2000). Synthesis and characterization of elastin-mimetic protein gels derived from a well-defined polypeptide precursor. *Macromolecules* **33,** 4809–4821.

Milewicz, D. M., Urban, Z., and Boyd, C. (2000). Genetic disorders of the elastic fiber system. *Matrix Biol.* **19,** 471–480.

Mithieux, S. M., Rasko, J. E., and Weiss, A. S. (2004). Synthetic elastin hydrogels derived from massive elastic assemblies of self-organized human protein monomers. *Biomaterials* **25,** 4921–4927.

Miyata, T., Uragami, T., and Nakamae, K. (2002). Biomolecule-sensitive hydrogels. *Adv. Drug Deliv. Rev.* **54,** 79–98.

Mongiat, M., Mungiguerra, G., Bot, S., Mucignat, M. T., Giacomello, E., Doliana, R., and Colombatti, A. (2000). Self-assembly and supramolecular organization of EMILIN. *J. Biol. Chem.* **275,** 25471–25480.

Muiznieks, L. D., Jensen, S. A., and Weiss, A. S. (2003). Structural changes and facilitated association of tropoelastin. *Archi. Biochem. Biophys.* **410,** 317–323.

Nakamura, T., Lozano, P. R., Ikeda, Y., Iwanaga, Y., Hinek, A., Minamisawa, S., Cheng, C. F., Kobuke, K., Dalton, N., Takada, Y., Tashiro, K., Ross Jr, J., Honjo, T., and Chien, K. R. (2002). Fibulin-5/DANCE is essential for elastogenesis *in vivo*. *Nature* **415,** 171–175.

Nath, N., and Chilkoti, A. (2002). Creating ''smart'' surfaces using stimuli responsive polymers *Adv. Mat. Proc.* **14,** 1243–1247.

Nicol, A., Gowda, D. C., and Urry, D. W. (1992). Cell adhesion and growth on synthetic elastomeric matrices containing Arg-Gly-Asp-Ser-3. *J. Biomed. Mat. Res.* **26,** 393–413.

Niklason, L. E., Gao, J., Abbott, W. M., Hirschi, K. K., Houser, S., Marini, R., and Langer, R. (1999). Functional arteries grown *in vitro*. *Science* **284,** 489–493.

Noblesse, E., Cenizo, V., Bouez, C., Borel, A., Gleyzal, C., Peyrol, S., Jacob, M.-P., Sommer, P., and Damour, O. (2004). Lysyl oxidase-like and lysyl oxidase are present in the dermis and epidermis of a skin equivalent and in human skin and are associated to elastic fibers. *J. Invest. Dermatol.* **122,** 621–630.

Ooyama, T., Fukuda, K., Oda, H., Nakamura, H., and Hikita, Y. (1987). Substratum-bound elastin peptide inhibits aortic smooth muscle cell migration *in vitro*. *Arteriosclerosis* **7,** 593–598.

Osakabe, T., Hayashi, M., Hasegawa, K., Okuaki, T., Ritty, T. M., Mecham, R. P., Wachi, H., and Seyama, Y. (2001). Age- and gender-related changes in ligament components. *Ann. Clin. Biochem.* **38,** 527–532.

Panitch, A., Yamaoka, T., Fournier, M. J., Mason, T. L., and Tirrell, D. A. (1999). Design and biosynthesis of elastin-like artificial extracellular matrix proteins containing periodically spaced fibronectin CS5 domains. *Macromolecules* **32,** 1701–1703.

Parks, W. C., and Deak, S. B. (1990). Tropoelastin heterogeneity: Implications for protein function and disease. *Am. J. Resp. Cell Mol. Biol.* **2,** 399–406.

Parks, W. C. (1997). Posttranscriptional regulation of lung elastin production. *Am. J. Respir. Cell Mol. Biol.* **17,** 1–2.

Partridge, S. M. (1962). Elastin. *Adv. Protein Chem.* **17,** 227–302.

Pasquali-Ronchetti, I., Fornieri, C., Baccarani Contri, M., and Volpin, D. (1979). The ultrastructure of elastin revealed by freeze-fracture electron microscopy. *Micron.* **10,** 89–99.

Pasquali-Ronchetti, I., Alessandrini, A., Baccarani Contri, M., Fornieri, C., Mori, G., Quaglino, D., Jr, and Valdre, U. (1998). Study of elastic fiber organization by scanning force microscopy. *Matrix Biol.* **17,** 75–83.

Patterson, C. E., Gao, J., Rooney, A. P., and Davis, E. C. (2002). Genomic organization of mouse and human 65 kDa FK506-binding protein genes and evolution of the FKBP multigene family. *Genomics* **79,** 881–889.

Peppas, N. A., and Langer, R. (1994). New challenges in biomaterials. *Science* **263,** 1715–1719.

Petersen, E., Wagberg, F., and Angquist, K. A. (2002). Serum concentrations of elastin-derived peptides in patients with specific manifestations of atherosclerotic disease. *Eur. J. Vascul. Endovascul. Surg.* **24,** 440–444.

Piontkivska, H., Zhang, Y., Green, E. D., and Elnitski, L. (2004). Multi-species sequence comparison reveals dynamic evolution of the elastin gene that has involved purifying selection and lineage-specific insertions/deletions. *Biomed. Cent. Genom.* **5,** 31.

Privitera, S., Prody, C. A., Callahan, J. W., and Hinek, A. (1998). The 67-kDa enzymatically inactive alternatively spliced variant of beta-galactosidase is identical to the elastin/laminin-binding protein. *J. Biol. Chem.* **273,** 6319–6326.

Ramirez, F. (2000). Pathophysiology of the microfibril/elastic fiber system: Introduction. *Matrix Biol.* **19,** 455–456.

Reinboth, B., Hanssen, E., Cleary, E. G., and Gibson, M. A. (2002). Molecular interactions of biglycan and decorin with elastic fiber components: Biglycan forms a ternary complex with tropoelastin and microfibril-associated glycoprotein 1. *J. Biol. Chem.* **277,** 3950–3957.

Reinhardt, D. P., Sasaki, T., Dzamba, B. J., Keene, D. R., Chu, M. L., Gohring, W., Timpl, R., and Sakai, L. Y. (1996). Fibrillin-1 and fibulin-2 interact and are colocalized in some tissues. *J. Biol. Chem.* **271,** 19489–19496.

Reitamo, S., Remitz, A., Tamai, K., Ledo, I., and Uitto, J. (1994). Interleukin 10 up-regulates elastin gene expression *in vivo* and *in vitro* at the transcriptional level. *Biochem. J.* **302,** 331–333.

Rich, C. B., Fontanilla, M. R., Nugent, M., and Foster, J. A. (1999). Basic fibroblast growth factor decreases elastin gene transcription through an AP1/cAMP-response element hybrid site in the distal promoter. *J. Biol. Chem.* **274,** 33433–33439.

Roark, E. F., Keene, D. R., Haudenschild, C. C., Godyna, S., Little, C. D., and Argraves, W. S. (1995). The association of human fibulin-1 with elastic fibers: An immunohistological, ultrastructural, and RNA study. *J. Histochem. Cytochem.* **43,** 401–411.

Rock, M. J., Cain, S. A., Freeman, L. J., Morgan, A., Mellody, K., Marson, A., Shuttleworth, C. A., Weiss, A. S., and Kielty, C. M. (2004). Molecular basis of elastic fiber formation. *J. Biol. Chem.* **279,** 23748–23758.

Rodgers, U. R., and Weiss, A. S. (2004). Integrin alpha v beta 3 binds a unique non-RGD site near the C-terminus of human tropoelastin. *Biochimica* **86,** 173–178.

Rosenbloom, J., Bashir, M., Yeh, H., Ornstein-Goldstein, N., Fazio, M., Kahari, V. M., and Uitto, J. (1991). Regulation of elastin gene expression. *Ann. N. Y. Acad. Sci.* **624,** 116–136.

Rosenbloom, J., Abrams, W. R., and Mecham, R. (1993). Extracellular matrix 4: The elastic fiber. *Fed. Am. Soc.Exper. Biol. J.* **7,** 1208–1218.

Rovira, A., Amedee, J., Bareille, R., and Rabaud, M. (1996). Colonization of a calcium phosphate/elastin-solubilized peptide-collagen composite material by human osteoblasts. *Biomaterials* **17,** 1535–1540.

Sakai, L. Y., Keene, D. R., and Engvall, E. (1986). Fibrillin, a new 350-kD glycoprotein, is a component of extracellular microfibrils. *J. Cell Biol.* **103,** 2499–2509.

Sandberg, L. B., Weissman, N., and Gray, W. R. (1971). Structural features of tropoelastin related to the sites of cross-links in aortic elastin. *Biochemistry* **10,** 52–56.

Sandberg, L. B., Soskel, N. T., and Leslie, J. G. (1981). Elastin structure, biosynthesis, and relation to disease states. *N. Engl. J. Med.* **304,** 566–579.

Senior, R. M., Griffin, G. L., and Mecham, R. P. (1980). Chemotactic activity of elastin-derived peptides. *J. Clin. Invest.* **66,** 859–862.

Senior, R. M., Griffin, G. L., and Mecham, R. P. (1982). Chemotactic responses of fibroblasts to tropoelastin and elastin-derived peptides. *J. Clin. Invest* **70,** 614–618.

Serafini-Fracassini, A., Field, J. M., and Spina, M. (1976). The macromolecular organization of the elastin fibril. *J. Mol. Biol.* **100,** 73–84.

Sittinger, M., Bujia, J., Rotter, N., Reitzel, D., Minuth, W. W., and Burmester, G. R. (1996). Tissue engineering and autologous transplant formation: Practical approaches with resorbable biomaterials and new cell culture techniques. *Biomaterials* **17,** 237–242.

Starcher, B. C. (2000). Lung elastin and matrix. *Chest* **117,** 229S–234S.

Stock, U. A., Wiederschain, D., Kilroy, S. M., Shum-Tim, D., Khalil, P. N., Vacanti, J. P., Mayer, J. E., Jr, and Moses, M. A. (2001). Dynamics of extracellular matrix production and turnover in tissue engineered cardiovascular structures. *J. Cell. Biochem.* **81,** 220–228.

Swee, M. H., Parks, W. C., and Pierce, R. A. (1995). Developmental regulation of elastin production. Expression of tropoelastin pre-mRNA persists after down-regulation of steady-state mRNA levels. *J. Biol. Chem.* **270,** 14899–14906.

Szabo, Z., Levi-Minzi, S. A., Christiano, A. M., Struminger, C., Stoneking, M., Batzer, M. A., and Boyd, C. D. (1999). Sequential loss of two neighboring exons of the tropoelastin gene during primate evolution. *J. Mol. Evol.* **49,** 664–671.

Tamburro, A. M., Guantieri, V., Daga-Gordini, D., and Abatangelo, G. (1977). Conformational transitions of alpha-elastin. *Biochim. Biophys. Acta* **492,** 370–376.

Tamburro, A. M., Guantieri, V., Scopa, A., and Drabble, J. M. (1991). Polypeptide models of elastin: CD and NMR studies on synthetic poly(X-Gly-Gly). *Chirality* **3,** 318–323.

Toonkool, P., Regan, D. G., Kuchel, P. W., Morris, M. B., and Weiss, A. S. (2001a). Thermodynamic and hydrodynamic properties of human tropoelastin. Analytical ultracentrifuge and pulsed field-gradient spin-echo NMR studies. *J. Biol. Chem.* **276,** 28042–28050.

Toonkool, P., Jensen, S. A., Maxwell, A. L., and Weiss, A. S. (2001b). Hydrophobic domains of human tropoelastin interact in a context-dependent manner. *J. Biol. Chem.* **276,** 44575–44580.

Uitto, J., Christiano, A. M., Kahari, V. M., Bashir, M. M., and Rosenbloom, J. (1991). Molecular biology and pathology of human elastin. *Biochem. Soc. Trans.* **19,** 824–829.

Umeda, H., Takeuchi, M., and Suyama, K. (2001). Two new elastin cross-links having pyridine skeleton. Implication of ammonia in elastin cross-linking *in vivo*. *J. Biol. Chem.* **276,** 12579–12587.

Urban, Z., and Boyd, C. D. (2000). Elastic-fiber pathologies: Primary defects in assembly and secondary disorders in transport and delivery. *Am. J. Hum. Genet.* **67,** 4–7.

Urry, D. W., Starcher, B., and Partridge, S. M. (1969). Coacervation of solubilized elastin effects a notable conformational change. *Nature* **222,** 795–796.

Urry, D. W., Cunningham, W. D., and Ohnishi, T. (1974). Studies on the conformation and interactions of elastin. Proton magnetic resonance of the repeating pentapeptide. *Biochemistry* **13,** 609–616.

Urry, D. W., and Long, M. M. (1977). On the conformation, coacervation and function of polymeric models of elastin. *Adv. Exp. Med. Biol.* **79,** 685–714.

Urry, D. W. (1978). Molecular perspectives of vascular wall structure and disease: The elastic component. *Perspect. Biol. Med.* **21,** 265–295.

Urry, D. W., Pattanaik, A., Xu, J., Woods, T. C., McPherson, D. T., and Parker, T. M. (1998). Elastic protein-based polymers in soft tissue augmentation and generation. *J. Biomat. Sci. Polymer Ed.* **9,** 1015–1048.

Urry, D. W., Hugel, T., Seitz, M., Gaub, H. E., Sheiba, L., Dea, J., Xu, J., and Parker, T. (2002). Elastin: A representative ideal protein elastomer. *Philos. Trans. R. Soc. London, Ser. B* **357,** 169–184.

van Hest, J. C. M., and Tirrell, D. A. (2001). Protein-based materials, toward a new level of structural control. *Chem. Comm.* **19,** 1897–1904.

Vardaxis, N. J., Boon, M. E., and Ruijgrok, J. M. (1996). Calcification of cross-linked collagen-elastin membrane implants *in vivo* and their proposed use in bone regeneration. *Biomaterials* **17,** 1489–1497.

Varga, Z., Jacob, M. P., Robert, L., and Fulop, T., Jr (1989). Identification and signal transduction mechanism of elastin peptide receptor in human leukocytes. *FEBS Lett.* **258,** 5–8.

Vrhovski, B., Jensen, S. A., and Weiss, A. S. (1997). Coacervation characteristics of recombinant human tropoelastin. *Eur. J. Biochem.* **250,** 92–98.

Vrhovski, B., and Weiss, A. S. (1998). Biochemistry of tropoelastin. *Eur. J. Biochem.* **258,** 1–18.

Weis-Fogh, T., and Anderson, S. O. (1970). New molecular model for the long-range elasticity of elastin. *Nature* **227,** 718–721.

Wendel, D. P., Taylor, D. G., Albertine, K. H., Keating, M. T., and Li, D. Y. (2000). Impaired distal airway development in mice lacking elastin. *Am. J. Resp. Cell Mol. Biol.* **23,** 320–326.

Wolfe, B. L., Rich, C. B., Goud, H. D., Terpstra, A. J., Bashir, M., Rosenbloom, J., Sonenshein, G. E., and Foster, J. A. (1993). Insulin-like growth factor-I regulates transcription of the elastin gene. *J. Biol. Chem.* **268,** 12418–12426.

Wu, W. J., Vrhovski, B., and Weiss, A. S. (1999). Glycosaminoglycans mediate the coacervation of human tropoelastin through dominant charge interactions involving lysine side chains. *J. Biol. Chem.* **274,** 21719–21724.

Yanagisawa, H., Davis, E. C., Starcher, B. C., Ouchi, T., Yanagisawa, M., Richardson, J. A., and Olson, E. N. (2002). Fibulin-5 is an elastin-binding protein essential for elastic fiber development *in vivo*. *Nature* **415,** 168–171.

Zhang, H., Apfelroth, S. D., Hu, W., Davis, E. C., Sanguineti, C., Bonadio, J., Mecham, R. P., and Ramirez, F. (1994). Structure and expression of fibrillin-2, a novel microfibrillar component preferentially located in elastic matrices. *J. Cell Biol.* **124,** 855–863.

Zhang, M., Pierce, R. A., Wachi, H., Mecham, R. P., and Parks, W. C. (1999a). An open reading frame element mediates posttranscriptional regulation of tropoelastin and responsiveness to transforming growth factor β1. *Mol. Cell. Biol.* **19,** 7314–7326.

Zhang, M. C., He, L., Giro, M., Yong, S. L., Tiller, G. E., and Davidson, J. M. (1999b). Cutis laxa arising from frameshift mutations in exon 30 of the elastin gene (ELN). *J. Biol. Chem.* **274,** 981–986.

AUTHOR INDEX

A

Abatangelo, G., 447
Abbott, W. M., 448
Abbs, S., 229
Abe, H., 353
Abe, M., 171
Aberer, W., 154
Abifadel, M., 429
Abrahamsberg, C., 156, 163
Abrams, W. R., 415, 416, 440, 441, 448
Adachi, E., 324, 344, 358
Adair, B. D., 63
Ades, L., 427
Aebi, U., 59, 84, 99, 114, 115, 116, 120, 121, 123, 124, 125, 126, 127, 128, 131, 132, 135, 144
Agre, P., 218, 230
Aguirre, L., 222
Aho, S., 157, 158
Aigner, T., 343
Aït-Ali, A., 447
Aizawa, M., 452
Akey, D. L., 92
Akhtar, N., 211
Akiyama, M., 416
Akiyama, T., 429
Akkari, P. A., 173
Ala-Kokko, L., 306, 326, 327, 357, 427
al-Alawi, N., 169
Alber, T., 19, 20, 43, 44, 47, 55, 56, 58, 60, 61, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 100, 105
Albertine, K. H., 448
Alberts, I. L., 354
Alcalai, R., 152
Aldridge, J., 228
Alessandrini, A., 356, 447, 448
Alfadhli, A., 59
Alix, A. J. P., 447
Allard, D., 429
Allen, P. G., 231, 232
Alloisio, N., 230
Almo, S. C., 214, 215
Altherr, M. R., 221
Amagai, M., 19
Amann, K. J., 209
Amano, S., 154
Amargo, E. V., 152
Amedee, J., 451
Ameen, N. A., 175, 177
Amenitsch, H., 366
Amigorena, S., 177
Amin, P., 331, 332
Amin, S. Z., 382
Amis, E. J., 273
Amrani, D. L., 260, 267
An, X., 221
Anand, R., 212
Anastasiadis, P. Z., 159, 160
Ancker, U., 227, 228
Anderson, J. M., 232
Anderson, N., 442
Anderson, S., 344
Anderson, S. O., 449
Andrä, K., 155, 156
Andreadis, A., 179
Andreoli, J., 170
Andres, A. C., 171
Andreucci, M., 162
Aneskievich, B. J., 222
Angelini, C., 229
Angles-Cano, E., 277
Angquist, K. A., 439
Angst, B. D., 149, 150
Anhalt, G. J., 158
Antin, P. B., 173
Aoki, N., 276
Apfelroth, S. D., 406, 408, 439
Apigo, J., 330

Appleton, R. E., 229
Aprison, B. S., 260
Apweiler, R., 226
Araya, E., 59
Archidiacono, N., 228
Arcovito, G., 269
Ardemasova, Z. A., 265
Arena, J., 363
Argraves, W. S., 274, 440
Ariëns, R. A. S., 250, 270, 271, 272, 282
Arking, D. E., 429
Armbrust, R., 230
Armstrong, D. K. B., 152
Arnaout, M. A., 162
Arndt, K. M., 58
Arner, A., 166
Arnott, S., 44, 303, 304, 329
Arnould, T., 161, 162
Arocha-Piñango, C. L., 280
Artlieb, U., 154, 156
Arvai, A. S., 63
Arvan, D. S., 262
Asada, S., 327
Asakura, S., 280
Asano, A., 216
Asante, E. A., 219
Ashburner, M., 226
Ashman, L. K., 415
Ashworth, J., 406, 426
Ashworth, J. L., 406, 411, 412, 413, 419, 420, 426, 428
Ashworth, J. N., 250
Asimos, A., 218, 230
Aspberg, A., 416, 440
Assmann, K. J., 213
Astbury, W. T., 248
Aszodi, A., 358, 360
Ata, K., 230
Atarashi, K., 304
Atkins, G. J., 416
Atkinson, R. A., 224
Aukerman, L., 451
Aumailley, M., 163, 382
Avolio, L., 153
Avruch, J., 177
Avsiuk, A., 178
Ayad, S., 382, 386, 387, 388
Ayala, Y. M., 263
Azhar, S., 176

B

Babinet, C., 163
Bacallao, R., 162
Baccarani-Contri, M., 416, 447, 448
Bach, T. L., 279
Bachinger, H. P., 57, 59, 319, 321, 325, 326, 377, 409, 410, 411, 413
Backman, L., 208, 223
Bacon-Baguley, T., 274
Badier-Commander, C., 451
Bailey, A. I., 221
Bailey, A. J., 346, 356, 453
Bailey, K., 248
Baines, A. J., 168, 206, 220
Bairoch, A., 13
Bajzar, L., 276
Baker, D. A., 53, 81
Baker, E., 416
Baker, N. E., 284
Bakker, M. A., 213
Baldeschwieler, J. D., 350
Baldock, C., 351, 354, 381, 382, 406, 414, 417, 418, 419, 420, 422, 423, 426
Bales, M. A., 206, 208
Ball, S., 382
Ball, S. G., 411, 420
Ballas, S. K., 229, 230
Ballester, F., 379
Baltimore, D., 224, 226
Bang, M. L., 173
Bann, J. G., 319
Bannon, L. J., 152
Bansal, M., 321
Banuelos, S., 214, 219, 221
Baradet, T. C., 268
Bardsley, W. G., 151, 161
Bareille, R., 451
Baribault, H., 174, 177
Barklis, E., 59
Barlow, P. N., 222
Barnes, G. R. G., 360
Barnes, M. J., 308, 309, 312, 318, 326, 327, 328, 329
Baron, M. D., 204
Barondes, S. H., 448
Bartlett, R., 228
Bartolo, C., 228
Bartunik, H. D., 379
Bar-Zvi, D., 204, 207, 210

Baschirotto, C., 153
Baselt, D. R., 350
Bashir, M. M., 416, 440, 441, 442, 443
Bassi, F. A., 269
Basso, C., 152
Bassuk, J. A., 168
Batchelder, L. S., 307, 316, 318
Bateman, A., 226
Bätge, B., 406, 411, 412, 413, 416
Batzer, M. A., 440
Bauce, B., 152
Bauer, C., 152, 153, 161
Baum, J., 312, 318, 322, 330, 331
Baumann, R. E., 282
Baumeister, W., 45, 51, 53, 54, 55, 60
Baumhueter, S., 261
Bautista, D. L., 95, 101
Bax, D. V., 407, 410, 412
Bdeir, K., 276
Beasley, J. R., 81
Beck, K., 57, 58, 331
Becker, B., 154, 156, 166, 180
Becker, G., 149, 154
Becker, K., 229
Becker, P. S., 230, 231
Beckmann, I., 170
Bedford, M. T., 224, 226, 227
Beffagna, G., 152
Beggs, A. H., 173, 228, 231, 232
Beguin, S., 274
Bell, W. R., 256, 259
Bella, J., 308, 309, 312, 314, 315, 316, 318, 321, 323, 330, 382
Bellamy, M., 168
Bellin, R. M., 114, 166, 180, 181
Bellingham, C. M., 439, 447, 451, 453
Bement, W. M., 180
Ben Abdeladhim, A., 230
Benazzoug, Y., 451
Bendeck, M. P., 387
Bennett, G. M., 272
Bennett, H., 210
Bennett, J. S., 277, 278
Bennett, M. J., 215
Bennett, V., 149, 211, 218, 220, 229, 230
Benoit, M., 377
Benson, M. A., 165
Bentley, D. R., 99, 212
Benyamin, Y., 216
Ben-Ze'ev, A., 159
Benzing, T., 161, 162
Berden, J. H., 213
Berendsen, H. J. C., 314
Berg, R., 19
Berger, B., 47
Berger, J. M., 47
Berisio, R., 309, 314, 316, 319, 322
Berjot, M., 447
Berland, P., 317, 319
Berliner, S. A., 278
Berman, H. M., 308, 309, 312, 314, 315, 316, 318, 321, 323, 330
Bernard, S. E., 407, 410, 412
Bernengo, J.-C., 344
Bernier, G., 152
Bernocco, S., 269, 317, 319, 344
Bernstorff, S., 366
Beroud, C., 427
Bershadsky, A. D., 163
Bertelson, C. J., 212, 228
Berthet-Colominas, C., 85, 344, 348, 350
Berthier, C., 59
Bertini, E., 376
Betre, H., 451
Bettecken, T., 228
Betts, L., 258, 259, 269
Bhate, M., 311
Bhatnagar, R. S., 321
Bhumralkar, M., 302
Biery, N. J., 428
Bies, R. D., 229
Bigi, A., 353
Bilak, M. M., 114
Bilak, S. R., 114
Bilgicer, B., 89
Bingler, M., 376
Binnie, C. G., 264
Binz, H. K., 59
Birk, D. E., 344, 358, 360
Birkenberger, L., 134
Birney, E., 226
Birtles, M. J., 360
Birukov, K. G., 448
Bisher, M. E., 127
Bishop, P. D., 270
Bishop, P. N., 357
Bittner, R., 156, 213
Bjorck, L., 306, 323
Bjork, J., 223
Black, D., 440

Blake, D. J., 165, 180, 210, 212, 213, 227
Blanchard, A., 210, 211, 223
Blaschke, U. K., 345
Blobel, G., 125, 168
Bloch, R. J., 213
Blombäck, B., 250, 260, 267, 269, 272, 273
Blonden, L. A. J., 228
Boak, B. B., 438, 448
Bobkov, A., 155, 215, 217
Bobrow, M., 212, 228
Bochicchio, B., 447
Bockenstedt, P., 274
Bocquet, J., 388
Bode, W., 263, 379
Boelens, R., 54
Boffa, M. B., 276
Bogin, O., 386, 387
Boguski, M. S., 67, 226
Bohn, W., 155
Boileau, C., 427, 429
Bokma, E., 67
Boles, J. O., 317, 319, 326
Bolliger, B., 267
Bolliger-Stucki, B., 259
Bolton, S. J., 214, 222
Bonadio, J., 406, 408, 439
Bonaglia, M. C., 153
Bondar, O., 379
Bondon, A., 221
Bonet-Kerrache, A., 222
Bonewald, L. F., 415
Bonilla, E., 213
Bonne, S., 151, 160
Bonzon, N., 453
Booms, P., 406, 409, 426, 427
Boon, M. E., 451
Boot-Hanford, R. P., 382, 386, 387, 388
Bordoli, L., 13
Borel, A., 439, 445
Borel, J. P., 388
Borgwardt, J. E., 151, 160
Borisy, G. G., 154, 155
Bornslaeger, E., 151
Bornslaeger, E. A., 150, 151, 160, 161
Bornslaeger, E. B., 150, 151
Borradori, L., 147, 150, 151, 152, 153, 154
Bortoletto, G., 221
Borza, D. B., 379
Boswell, B. A., 411
Bot, S., 440
Boudko, S., 309
Bouez, C., 439, 445
Bouhassira, E. E., 230
Bourenkov, G. P., 84, 126, 379
Bousquet, O., 168, 180
Boutillon, M.-M., 388
Bova, S., 153
Bowers, B., 157, 158
Bowers, E., 152, 161
Boxer, M., 427
Boyd, C. D., 440, 442, 450
Boyd, S. M., 58, 95
Boyer, M., 216
Brabec, V., 229, 230
Bradby, G. V., 355
Bradshaw, A. D., 359
Bradshaw, J. P., 348
Braga, S., 228
Bragheri, R., 153
Braile, D. M., 451, 452
Brain, M. C., 229
Brakebusch, C., 163
Bramham, J., 222
Brandes, D., 355
Brandin, E., 206, 210
Branson, H. R., 38, 39
Branton, D., 204, 206, 207, 210, 224, 225
Brass, A., 358
Braswell, E., 330
Brazel, D., 377
Bredt, D. S., 221
Bregant, T., 100
Brennan, S. O., 264, 281
Brenner, S., 25
Bresnick, A. R., 216
Bressan, G. M., 440, 448
Bretcher, L. E., 322
Breton-Gorius, J., 262
Bretscher, A., 172
Bretscher, L. E., 321, 322
Brettel, M., 124, 132, 160
Breuning, M. H., 162
Bridges, D., 176
Bridgman, P. C., 179
Brinckmann, J., 406, 411, 413
Brittingham, R. J., 324
Britton, J. E., 166
Brocheriou, V., 166
Broderick, M. J. F., 28, 215, 224

Brodsky, B., 306, 308, 309, 311, 312, 314, 315, 316, 318, 320, 321, 322, 323, 324, 325, 326, 327, 328, 330, 331, 332, 343, 346, 350, 353
Brodsky-Doyle, B., 348
Brody, T., 226
Broekelmann, T., 410, 412, 413, 414, 415, 416, 444, 445, 446
Brokstein, P., 226
Brooke, B., 438, 448
Broome, R. L., 174, 176
Brosius, M., 165
Brown, A., 152, 179
Brown, E. M., 349
Brown, J. C., 380
Brown, J. H., 21, 25, 51, 61, 124, 129, 255, 258, 270
Brown, N. H., 147
Brown, R. A., 56, 89
Brown-Augsberger, P., 415, 444, 445, 446
Bruckner, P., 323, 345, 346
Bruckner-Tuderman, L., 415, 416
Bruijin, J. A., 386, 387
Brunger, A. T., 48, 61
Brunmark, C., 379
Bryan, J. T., 59, 122, 124, 128, 129, 131
Bucher, P., 13
Buckley, K. M., 147, 149
Buczek-Thomas, J. A., 440
Budzynski, A. Z., 250, 264, 270
Buevich, A. V., 223
Bugge, T. H., 282
Bujia, J., 451
Bujnicki, J. M., 306
Bullitt, E., 253, 254, 257, 267
Bullough, P. A., 99
Bullrich, F., 154
Bulman, D. E., 219, 228
Bunce, L. A., 278
Bunton, T. E., 429
Burakoff, S. J., 225
Bürer, A., 125
Burgeson, R. E., 154, 344
Burghes, A. H., 228
Burgoon, M. P., 284
Burjanadze, T. V., 324
Burkhard, P., 20, 21, 44, 45, 48, 53, 59, 82, 84, 86, 99, 121, 123, 124, 125, 126, 127, 128, 129, 131, 132
Burman, D., 451
Burmeister, W. P., 60
Burmester, G. R., 451
Burridge, K., 167, 212, 225
Burrow, C. R., 162
Bursaux, E., 229
Busaan, J. L., 227
Busuttil, S. J., 279
Butkowsky, R. J., 379
Butler, R., 212
Byers, P. H., 329, 330, 331, 332
Byers, T. J., 204, 206, 207, 210
Byrne, A. M., 231
Byth, B. C., 213
Byun, Y., 174

C

Cabezas, J. C., 152, 161
Caen, J. P., 280
Cai, P., 445
Cain, S., 407, 414, 415, 422
Cain, S. A., 411, 414, 420, 444
Cain, W. J., 276
Cairns, L., 174
Califano, J. C., 213
Calkins, C. C., 151
Callahan, J., 444
Callahan, J. W., 444
Camacho Vanegas, O., 376
Cameron, G. J., 354
Campanelli, J. T., 226
Campbell, I. D., 408, 423
Campbell, K. M., 56, 60, 92, 105, 213
Campbell, K. P., 212, 213, 219, 227, 229
Campbell, L., 213
Campbell, S. S., 174
Candi, E., 135, 170
Cann, J. R., 349
Cantieri, J. S., 168, 169
Cao, S. N., 410, 412, 413
Capetanaki, Y., 174
Cappellini, D. M., 229
Cappello, J., 451
Caput, D., 125
Carden, M. J., 19
Cardinali, B., 269
Cardy, C. M., 408, 419, 423
Carl, P., 218
Carlsson, K., 269, 272, 273

Carlsson, L., 164
Carmical, S. G., 410, 412, 413
Caron, A., 174
Carpen, O., 212, 221
Carr, C. M., 84, 99
Carr, M. E., 272, 273
Carreras, I., 443
Carroll, H., 358
Carroll, W. R., 275
Cartaud, A., 161
Cartaud, J., 161
Carter, A. M., 282
Carugo, K. D., 214
Carvajal-Huerta, L., 152, 161
Casanova, M. L., 176
Cascio, D., 58, 93, 94, 101
Casella, J. F., 218, 230
Caskey, C. T., 229
Cassidy, K., 353
Castagnola, P., 387
Castano, J., 160
Castanon, M. J., 149, 154, 156
Castellani, I., 440, 448
Castillos, F. A., 114
Castresana, J., 206, 207, 211, 214, 221
Cataldo, A., 179
Caulin, C., 174
Cederholm-Williams, S. A., 250
Celniker, S. E., 226
Cenizo, V., 439, 445
Centner, T., 173
Cerasoli, E., 85, 97, 98, 100, 103
Cerritello, A., 153
Cerutti, L., 13
Cesareni, G., 226
Ceyhan, B., 100
Chaiken, I. M., 103
Chaikof, E. L., 451
Chajek-Shaul, T., 152
Chakravarti, S., 358
Chamberlain, J. S., 219, 227, 229
Champetier, S., 173
Chan, V. C., 331
Chan, W. K., 179
Chandrasekharan, R., 312, 314
Chang, L., 178, 179, 180
Chang, R., 174
Chanut-Delalande, H., 344
Chao, H. G., 95
Chao, H. M., 95, 101
Chapman, J. A., 346, 358, 388
Chapman, J. M., 277
Chapon, F., 174
Charbonneau, N. L., 406, 408, 428
Chase, B., 264
Chateau, D., 174
Chaudhary, N., 125
Chaudhry, C., 99
Chavanas, S., 163
Chelly, J., 228
Chen, C. B., 327, 328, 330
Chen, F., 174
Chen, J. D., 228
Chen, J. M., 349
Chen, K. J., 330
Chen, S. W., 210
Chen, X., 151, 152, 160
Chen, Y., 349
Chen, Y. Q., 443
Chen, Z. P., 172
Cheng, C.-F., 416, 440
Cheraud, Y., 166
Cheresh, D. A., 278, 412
Cherry, J. H., 452
Cherry, J. M., 226
Chervitz, S. A., 226
Chew, M. W. K., 18, 303, 389
Chien, K. R., 416, 440
Chilcote, R. R., 230
Child, A., 427
Childs, B., 350
Chilkoti, A., 451, 452
Chin, S. S. M., 134, 174
Ching, G. Y., 114
Chiou, S. S., 231
Chipev, C. C., 168
Chmielewski, J., 82
Choe, S., 58, 93, 94, 101
Choe, V., 343, 346
Choi, H. J., 150
Chong, N. H. V., 382
Chong, P. C. S., 92, 102
Chothia, C., 40, 45, 83
Chou, Y.-H., 19, 114, 132, 134, 154, 172, 181
Choulier, L., 408, 409, 412, 418, 420, 423, 425
Christiano, A. M., 154, 440, 442, 443
Christiansen, D. L., 349
Christmann, D., 353
Chritchley, D. R., 216

Chrzanowska-Wodnicka, M., 225
Chu, C. L., 440
Chu, D., 172
Chu, M.-L., 149, 153, 323, 345, 376, 416, 440
Chung, D. W., 257, 259, 260, 261, 277
Chung, L., 319
Chung, S. I., 170
Chung, W., 226
Ciani, B., 100
Cicchetti, F., 416
Cicchetti, P., 226
Cierniewski, C. S., 270
Ciliberto, C. H., 444
Cines, D. B., 276
Clark, A. G., 206, 208
Clark, J. I., 174
Clark, K. L., 224, 226
Clarke, A. W., 444
Claudepierre, T., 213
Claustres, M., 429
Cleary, E. G., 414, 415, 439, 440
Cleveland, D. W., 114, 128, 144, 152, 153, 180, 181
Clos, J., 99
Clubb, B. H., 154
Cluzel, C., 306
Coates, D., 226
Cocco, M. J., 59
Cochran, A. G., 47
Coetzer, T. L., 229
Coffey, A. J., 212
Cogne, S., 327
Cohen, C., 14, 21, 25, 44, 45, 48, 51, 53, 60, 61, 81, 124, 129, 253, 254, 255, 257, 258, 259, 265, 267, 270, 272
Cohen, C. M., 167
Cohen, F. E., 70
Cohen, G. M., 174
Cohn, E. J., 250
Cole, A. A., 349
Cole, C. G., 212
Cole, N. M., 219
Cole, W. G., 329, 330, 331
Colella, S., 278
Collen, D., 275
Collet, J.-P., 275, 277, 280
Colletti, C. A., 228
Collichio, F. A., 262
Collins, F. S., 230
Collod-Beroud, G., 427, 429
Colman, D. R., 134
Colombatti, A., 440
Colucci-Guyon, E., 163, 177
Comeglio, P., 427
Comella, N., 162
Comelli, A., 221
Common, J., 152, 161
Compton, J. G., 117, 133, 135
Comtois, A., 153
Condeelis, J., 210, 216
Conn, K. J., 443
Conticello, V. P., 99, 451
Conway, J. F., 47, 57, 86, 87, 88, 89, 90, 91, 92, 96, 117, 124, 125, 127, 129, 360
Cooke, P., 171
Coovert, D. D., 228
Corbascio, G. C., 278
Corcoran, C. M., 150, 151, 160
Cordingley, H. C., 151, 161
Corey, R. B., 38, 39, 55
Corradin, G., 96, 98
Corrado, K., 229
Correia, I., 172
Correia, L. A., 231, 232
Corson, G. M., 406, 408, 411, 428
Cosgroves, D., 379
Cote, H. C., 257, 259, 280
Couchman, J. R., 152
Coulombe, P. A., 144, 168, 173, 174, 177, 180
Coupal, E., 19, 204
Courtney, M. A., 262
Courtois, G., 261
Courvalin, J. C., 161
Couture-Tose, E., 306
Covone, A. E., 228
Cowan, P. M., 302, 307, 350
Cox, G. A., 229
Crabtree, G. R., 260, 261
Craig, A. S., 350, 353, 360
Craig, L., 63
Craig, R., 215, 217
Craig, S. W., 164, 165, 213
Cramer, E. M., 262
Cravchik, A., 226
Craven, N. M., 426
Cretoiu, J. S., 410, 412, 413
Crewther, W. G., 14, 23, 58, 126, 127, 128
Crick, F. H. C., 18, 38, 39, 40, 302, 307, 308, 311
Crissman, J. W., 451

Crissman, M., 451
Critchley, D. R., 166, 181, 204, 207, 210, 211, 214, 222, 223
Cross, R. A., 172
Crossin, K. L., 284
Crouch, E., 416
Crowe, C. A., 376
Cryns, V. L., 174
Cumiskey, A. M., 412
Cummings, C., 351, 388, 417
Cuniberti, C., 269
Cunliffe, S., 426
Cunningham, B. A., 284
Cunningham, W. D., 445, 447, 450
Curran, J., 60
Curtis, P., 19
Curtis, P. J., 204
Cusack, S., 85
Cutillo, S., 229

D

Daculsi, G., 453
Daga-Gordini, D., 440, 447
Dagonnet, F. B., 280
Dahlbäck, B., 416, 439
Dahlbäck, K., 416, 439
Daigle, N., 173, 175, 177
Dale, B. A., 168, 169, 170
D'Alessio, M., 406
Dalhaimer, P., 218
Dallas, S. L., 415
Dalla Vecchia, F., 221
Dalpe, G., 153
Dalton, N., 416, 440
Damonte, G., 269
D'Andrea, L., 59
Dang, C. V., 256, 259
Dang, Q. D., 263
Danieli, G. A., 152
Daniell, H., 452
Danielson, K. G., 357, 358
Dano, K., 282
Danowski, B. A., 164, 213
Dantas, G., 81
Danton, M. J., 282
Dardik, B. N., 265, 267, 280
Darling, J. M., 174
Darras, B. T., 228
Dasgupta, J., 359
Daub, E., 257
Daugherty, C. C., 281, 282
Daugherty, M. A., 270
David, C. N., 377
Davidson, J. M., 282, 359, 440, 442, 443
Davie, E. W., 257, 259, 260, 261, 277
Davie, J. V., 229
Davies, K. E., 165, 166, 209, 210, 212, 213, 214, 219, 220, 222, 227, 228
Davis, E. C., 406, 408, 412, 416, 417, 419, 438, 439, 440, 443, 448
Davis, J. M., 326
Davison, M. D., 204, 207
Davour, O., 439, 445
Day, K., 283
Dea, J., 449
Deak, S. B., 440, 442
Debelle, L., 438, 442, 447, 450
Debili, N., 262
de Boer, J. P., 276
De Castro, E., 13
Decker, S., 416
De Cristofaro, R., 263
Degen, J. L., 279, 281, 282
Degenstein, L., 152, 161, 163
Degrado, N. F., 81
Degrado, W. F., 58, 82, 93, 94, 101
de Groot, P. G., 278
Degterev, A., 443
de Heer, E., 162, 386, 387
Deillon, C., 56, 88
Deininger, K., 224, 226, 227
Deisenhofer, J., 29, 30
Dejana, E., 278
de la Chapelle, A., 173, 228
Delacoux, F., 327
Delaunay, J., 229, 230
Delorenzi, M., 47
Delouvee, A., 163
de Nechaud, B., 174
Deng, H., 224
Denk, H., 154, 156, 174
Denker, B. M., 232
Denning, M. F., 151
Dente, L., 226
Denton, M., 228
De Paepe, A., 427
de Pereda, J. M., 155, 215, 217
Deprez, P., 306, 326, 327, 328

Derancourt, J., 167
De Repentigny, Y., 153
De Rider, M. L., 322
DeRosier, D., 215, 216
Derwin, K. A., 358
De Spirito, M., 269
Detken, A., 100
Deutzmann, R., 323, 379, 380
DeWolf, C., 221
Dewson, G., 174
De Zenzo, G., 226
Dhaliwal, A. S., 151
Dhall, D. P., 273
Dhermy, D., 230
Diamond, A. M., 356
Diaz, L. A., 161
Diaz-Ricart, M., 274
Dibello, P. M., 264
Di Cera, E., 263, 272
Dickerson, S. K., 259
Dickson, G., 213, 219
DiColandrea, T., 157, 158
Dietz, H. C., 428, 429, 439
Dietzschold, B., 19
DiGiovanna, J. J., 135
Dijkman, H. B., 213
Di Lullo, G. A., 326, 327
Dinakarpandian, D., 319
Dinauer, P. A., 256
Dinsdale, D., 174
DiOreo, J. P., 270
DiOrio, J. P., 267, 268
Discher, D. E., 218
Di Stasio, E., 272
Dixit, V. M., 274
Dixon, T. W., 19, 58
Djabali, K., 174
Djinovic-Carugo, K., 217, 220, 221
Dmytrenko, G. M., 213
Dobson, C. M., 100
Dogic, D., 163
Doi, M., 322
Doi, T., 306, 327
Doliana, R., 440
Dominguez, R., 25, 61, 255, 258
Domon, T., 363
Donner, K., 173
Donovan, J. W., 260
Donovan, P. J., 357, 358
Dooley, T. P., 149, 157
Doolittle, L. R., 256, 279
Doolittle, R. F., 61, 250, 253, 255, 256, 257, 258, 259, 260, 264, 267, 269, 270, 271, 276, 283
Dordick, J. S., 452
Doss-Pepe, E., 306, 326, 327, 328
Dours-Zimmerman, M., 416
Dover, S. D., 44
Dowling, J., 152, 153, 163
Dowling, L. M., 23, 126, 128
Downing, A. K., 406, 408, 409, 410, 423
Downing, J. S., 114
Downing, S. W., 114
Doyle, B. B., 356
Drabble, J. M., 450
Dreher, M. R., 451
Dreisewerd, K., 343
Drenckhahn, D., 155, 231, 232
Drew, A. F., 282
Dreyfus, H., 213
Drickamer, K., 327, 328, 330
Drummond, I., 161, 162
Du, X., 279
Duance, V. C., 343, 419
Dube, M., 232
Dublet, B., 387
Dubowitz, V., 173
Dubreuil, R. R., 204, 206, 207, 210
Ducluzeau, M. T., 230
Duggan, D. J., 228
Dunach, M., 160
Dupret, J. M., 174
Durbin, R., 226
Dure, L. III, 53
Dusek, R. L., 147, 160
Dutoya, S., 448, 451
Dwyer, M. A., 80, 81
Dyr, J. E., 260
Dzamba, B. J., 416, 440

E

Eady, R. A. J., 152, 154
Earnest, J. P., 213
Earnshaw, W. C., 174
Eaton, M., 309, 314, 321, 323, 330
Ebashi, S., 211
Eber, S. W., 229, 230
Eble, J. A., 327

Eck, M. J., 210, 224, 226
Eckardt, C., 382
Eckardt, U., 382
Eckert, D. M., 92
Eckes, B., 163
Eddy, S. R., 226
Edelberg, J. M., 277
Edelman, G. M., 284
Edlund, M., 211
Edlund, U., 223
Edman, A. C., 448
Edwards, Y. H., 213
Egelman, E. H., 215, 217
Eger, A., 154
Ehmann, C., 170, 171
Ehmer, S., 213
Ehrlich, H. J., 276
Eichenberger, D., 363
Eichwald, E., 438, 448
Eikenberry, E. F., 345, 350, 353
Einarson, M. B., 163
Eisenberg, D., 58, 93, 94, 101, 215
Ekani-Nkodo, A., 356
Elgart, G. W., 149
Elias, P. M., 135
Elliott, B., 254, 257
Elliott, C. E., 154, 156
Ellis, J. M., 229
Elnitski, L., 440, 442
Elshourbagy, N., 67
Elson, E. L., 217
Emery, A. E., 227
Emsley, J., 308, 309, 312, 318, 328, 329
Enders, G. C., 379
Endo, M., 211
Endo, T., 222
Endow, S. A., 99
Engel, A. M., 51, 53, 60
Engel, J., 21, 54, 58, 59, 84, 124, 309, 314, 322, 323, 377, 380, 382, 386
Engel, U., 377
Engelhardt, H., 53, 60
England, S. B., 219, 228
Englander, S. W., 312
Engle, L. J., 179
Engvall, E., 416, 439
Eramaa, M., 204, 209
Erickson, H. P., 149, 253, 265, 267, 284
Eriksen, T. A., 419
Eriksson, J. E., 175, 177
Ervasti, J. M., 164, 166, 209, 213, 215, 217, 220
Eshkind, L. G., 151, 174
Espanel, X., 224, 226, 227
Esquivel, C. O., 174
Evans, E. A., 218
Evans, R. M., 177
Evdokiou, A., 416
Everse, S. J., 255, 257, 258, 259, 260, 269, 270, 276, 283
Exposito, J. Y., 306
Eyre, D. R., 346, 357, 382
Ezra, E., 382, 383

F

Fabbrizio, E., 213, 222
Faigle, W., 177
Fairbrother, U., 213
Fairley, J. A., 161
Fairman, R., 95
Faix, J., 59, 84
Falcone, J. C., 231
Falls, L. A., 270
Falquet, L., 13
Fan, K., 80
Fan, M. J., 161
Fan, P., 312, 318, 322
Fanning, J. C., 415, 422, 439
Fantner, G. E., 356
Faraco, J., 416
Fardeau, M., 173, 174
Farjanel, J., 363
Farndale, R. W., 308, 309, 312, 318, 326, 327, 328, 329
Farquhar, M. G., 232
Farrell, D. H., 264, 270, 277, 278
Fashena, S. J., 163
Fassler, R., 156, 163, 358, 360
Fatah, K., 269, 272, 273
Fattoum, S., 230
Faulkner, G., 221
Faulkner, J. A., 219
Faundez, V., 178
Faure, A., 174
Faury, G., 448, 451
Fauser, C., 377
Faust, L. Z., 58
Favre, B., 150, 151, 152, 153, 154

Fazio, M. J., 440, 441
Fedorov, A. A., 214, 215
Feener, C., 212
Fehrentz, J. A., 213
Feja, B., 132
Fekete, A., 441
Feldman, L., 451
Fellowes, A. P., 281
Feng, D., 388, 389
Fenselau, C. C., 256
Fenwick, R., 229
Feo, C., 230
Ferber, A., 279
Ferguson, C., 221
Ferguson, K. M., 224, 225
Ferrari, F. A., 451
Ferrari, P., 448
Ferre, N., 453
Ferri, F., 269
Ferry, J. D., 272, 273
Fertala, A., 324
Fesik, S. W., 225
Fessler, L. I., 344, 358
Fichard, A., 327, 344
Fichera, A., 89
Ficheux, D., 307
Fichou, Y., 221
Fidzianska, A., 228
Field, J. M., 447
Fields, G. B., 302, 318, 319, 327, 328
Fietzek, P. P., 18, 346
Figgemeier, E., 21
Figueroa, Y., 175, 177
Fillers, J. P., 25, 60, 257
Findlay, D. M., 416
Finlay, L., 59
Finn, R. D., 226
Finnis, M. L., 415
Fiorelli, G., 229
Fiori, S., 328, 329
Firth, E. C., 360
Fischbeck, K. H., 228
Fischer, I., 156, 159
Fischetti, R. F., 350, 352
Fisher, D. Z., 125
Fitch, J. M., 344, 358
Fite, D., 344
FitzGerald, O., 359
Flanagan, L. A., 179
Fleckman, P., 168, 169
Fleischmann, W., 226
Fleischmayer, R., 382
Fless, G., 277
Flick, J. D., 279
Flick, M. J., 282
Flint, T. J., 228
Flitney, F. W., 156, 163, 180
Flory, P. J., 449
Flynn, N. T., 101
Fogo, A., 231
Foisner, R., 144, 154, 155
Foley, J. E., 344
Folkhard, W., 353, 366
Fong, J. H., 58
Fontaine, V., 213, 451
Fontanilla, M. R., 443
Fontao, L., 150, 151, 152, 154, 155
Forest, K. T., 63
Forget, B. G., 19, 204, 209, 211, 220, 229, 230, 231
Fornace, A. J., 260
Fornieri, C., 447, 448
Forrer, P., 59
Forrest, S. M., 219, 228
Fortini, M. E., 226
Foster, J. A., 416, 440, 443, 449
Foster, J. L., 222
Fournier, C., 277
Fournier, M. J., 451, 452
Fowler, V. M., 172, 173
Fowler, W. E., 253, 265, 267
Fox, J. E., 213
Franc, S., 356
Franci, C., 160
Francis, C. W., 250, 252, 255, 259, 261, 262, 264, 274, 275, 278
Francke, U., 228, 230, 416
Frank, S., 59, 309
Franke, R. P., 231
Franke, W. W., 120, 149, 151, 160, 174, 177
Franzblau, C., 444
Frappier, T., 167
Fraser, R. D. B., 14, 18, 21, 23, 44, 58, 117, 121, 124, 126, 127, 132, 303, 307, 308, 314, 316, 319, 321, 322, 350, 351, 353
Frasson, M., 213
Fratzl, P., 351, 353, 356, 362, 365, 366
Fratzl-Zelman, N., 351
Frazier, W. A., 274
Freeman, L., 407, 414, 415, 422, 444

French, C. J., 446
Frenkiel, T. A., 224
Freyssinet, J. M., 265
Freyzon, Y., 58
Fried, M. G., 270
Friedman, D. B., 379
Frischmeyer, P. A., 429
Fu, H., 177
Fu, S. Y., 227
Fu, Y., 260
Fuchs, E., 144, 150, 151, 152, 153, 160, 161, 163, 181
Fuchs, J. R., 451
Fuchs, P., 156, 159, 163
Fujii, I., 103
Fujimaki, N., 156
Fujimoto, N., 416
Fujinaga, M., 85
Fujita, D. J., 175
Fujita, E., 175
Fujita, Y., 363
Fujiwara, S., 13, 17, 159
Fukami, K., 222
Fukuda, K., 448
Fukumi, K., 222
Fuller, G. M., 262
Fullwood, N. J., 360
Fulop, T., Jr., 448
Funahashi, T., 304
Fung, Y. C., 446
Furst, D. O., 156, 165
Furthmayr, H., 380, 382, 386
Furuhashi, K., 222
Furukawa, Y., 429
Fygenson, D. K., 356

G

Gaboriau, F., 221
Gabor Miklos, G. L., 226
Gabriel, D. A., 272, 279
Gabrielian, A., 226
Gache, Y., 163
Gaffney, D. P., 269
Gagescu, R., 154
Gailit, J., 278
Galanakis, D. K., 262, 273, 278, 279, 281
Galkin, V. E., 215, 217
Galla, H. J., 345
Gallagher, P. G., 229, 230
Gallay, J., 221
Galle, R. F., 226
Galli, T., 177
Gallicano, G. I., 152
Gambee, J. E., 57, 410, 411, 413
Gan, S.-Q., 17, 117, 118, 133, 168
Gao, J., 166, 439, 443, 445, 448
Gao, Y. S., 177
Garcia, C. A., 229
Garcia-Alvarez, B., 155, 215, 217
Garcia de Herreros, A., 160
Garcia-Echeverria, C., 100
Gardner, K., 220
Gardner, R. J., 228
Garnier, S., 448
Garrod, D. R., 151, 161
Garrone, R., 306, 317, 319, 327, 344, 388
Gartner, T. K., 274
Garza, N., 317, 319, 326
Gattiker, A., 13
Gaub, H. E., 218, 221, 377, 449
Gaudry, C. A., 147, 160
Gautel, M., 217, 220, 221
Gayraud, B., 429
Geercken, E., 366
Geercken, W., 353
Geerts, D., 150, 151, 152, 153, 154, 155
Geiger, B., 163
Geisler, N., 123, 168
Gelbart, W. M., 226
Genet, S. A., 228
Geng, L., 162
Gengyo-Ando, K., 25
Gentz, R., 19
Georgatos, S. D., 114, 115, 167, 168
George, A., 348
George, P. M., 281
George, R. A., 226
Gerecke, D. R., 344, 387
Gerig, J. T., 307, 318
Gerisch, G., 59, 84
Getsios, S., 147, 152, 160
Getzoff, E. D., 63
Geubtner, J., 428
Ghanem, A., 230
Ghori, N., 174
Ghosh, I., 99, 101, 104
Ghosh, S., 174
Giacomello, E., 440

Giannelli, F., 212
Gibney, E., 344, 358
Gibson, C. W., 448
Gibson, M. A., 414, 415, 416, 422, 439, 440
Gibson, T., 224
Gibson, T. J., 207, 212, 215, 219
Giglio, S., 153
Gilbert, B., 151, 160
Gilbert, S., 173, 175, 177
Gilgenkrantz, H., 228
Gill, S. R., 128
Gillanders, E. M., 376
Gilligan, D. M., 211, 218, 229
Gilliland, G. L., 263
Gilman, P. B., 256, 282
Gilpin, C. J., 351, 354, 417, 420
Giltay, R., 416
Gimona, M., 165, 171, 172, 213, 214
Ginsberg, M. H., 278
Giorda, R., 153
Giordano, M., 346, 349
Giraud-Guille, M.-M., 363
Girbal-Neuhauser, E., 169
Giro, M. G., 442, 447
Gish, R., 174
Giusti, B., 376
Glagov, S., 448
Glanville, R. W., 323, 343, 382, 415, 417, 420
Glattauer, V., 327
Glenney, J. R., Jr., 172
Gleyzal, C., 439, 445
Glover, J. N., 57
Gobbi, P., 356
Godal, H. C., 259
Godsel, L. M., 147, 151, 152, 160
Godyna, S., 274, 440
Goetz, J., 330
Goh, M. C., 355, 356
Göhring, W., 416, 440
Goissis, G., 451, 452
Golbik, R., 53, 60, 327
Goldbaum, D. M., 256
Goldberg, L. R., 228
Goldfinger, L. E., 147, 152, 163
Goldie, K., 51, 53, 60, 219
Goldman, A., 154
Goldman, B., 154
Goldman, R. D., 19, 114, 132, 134, 144, 154, 156, 163, 172, 174, 178, 179, 180, 181
Goldsmith, L. A., 161
Goldsmith, S., 214, 215
Goldstein, B. S., 360
Goldstein, L. S., 179, 180, 204, 207, 210, 226
Golemis, E. A., 163
Golige, A., 363
Gong, F., 226
Gonzales, L., 89, 100
Gonzales, L., Jr., 56
Gonzales, M., 156, 163
Gonzalez, L., 89
Gonzalez, L. J., 88, 89, 92, 105
Gonzalez-Gronow, M., 277
Gonzalez-Vitale, G. C., 355
Goodman, M., 302
Goodship, A. E., 360
Gorbea, C., 30
Gordon, B. R., 277
Gordon, J. I., 67
Gordon, M. K., 344, 358, 387
Gorham, S. D., 356
Gorkun, O. V., 254, 258, 259, 264, 267, 269, 270
Goshima, F., 181
Gosline, J. M., 446, 449, 453
Gosling, S., 357
Gotte, L., 447
Goud, H. D., 443
Goulielmos, G., 114, 115
Gounari, F., 114, 115
Gowda, D. C., 452
Grady, R. M., 166
Granata, V., 316, 319
Grange, R. W., 166
Granger, B. L., 166, 168
Grant, G. A., 274
Grant, M. E., 303, 304, 382, 388, 417
Grant, P., 175
Grant, P. J., 250, 270, 271, 272, 282
Grassel, S., 323, 345
Gratzer, W. B., 204, 210, 221
Graves, D. J., 136
Gray, W. R., 445, 449
Greco, M., 269
Green, E. D., 212, 440, 442
Green, H., 157
Green, J. D., 100
Green, K. J., 146, 147, 149, 150, 151, 152, 153, 154, 158, 159, 160, 161, 174
Greenberg, C. R., 228

Greenberg, C. S., 250, 270, 271, 272, 282
Greenfield, N. J., 25, 61
Gregg, K., 16, 17
Gregorio, C. C., 172, 173
Gregory, K. W., 451, 452
Gregory, S. L., 147
Grieninger, G., 260, 261
Griffin, G. L., 448
Griffiths, C. E., 426
Grigera, J. R., 314
Grimbergen, J., 280
Grimes, J. M., 408, 409, 412, 418, 420, 423, 425
Grishin, N. V., 29, 80
Gronholm, M., 221
Gronska, M. A., 387
Groos, S., 382
Groot, K. R., 158
Gross, J., 355
Grosshans, E., 330
Grosso, L. E., 443
Groult, V., 448
Gruber, M., 55, 72, 82, 86
Gruenbaum, Y., 144
Grum, V. L., 218
Grund, C., 177
Gruning, W., 161
Grutter, M. G., 56, 59, 88
Grzesiek, S., 377
Gualtieri, R. J., 229
Guan, P., 226
Guantieri, V., 447, 450
Guda, C., 452
Guicheney, P., 166, 174
Guinto, E. R., 263
Gumbiner, B. M., 152, 159
Gundersen, G. G., 178
Gunwar, S., 379
Guo, A. W., 209
Guo, D. C., 427
Guo, L., 99, 152, 163
Guo, M., 275
Guo, X., 221
Gupta, A., 231
Gurewich, V., 277
Gurnon, D. G., 104, 105
Gurusiddappa, S., 317, 319, 326
Gusev, N. B., 171, 172
Gustavson, K. H., 317, 319
Gutsmann, T., 356
Gutte, B., 56, 88
Guyon, J. R., 165, 166
Guzei, I. A., 321
Gyoeva, F., 178

H

Haas, T. A., 279
Hachisuka, H., 156
Hada, M., 274
Haffner, C., 151, 160
Haftek, M., 169
Hagemeier, C., 409
Hagg, R., 345
Hahn, A. W. A., 448
Hahn, R., 344
Hai, J. W., 95
Hai, R., 58
Haidaris, P. J., 262
Haimovich, B., 222
Hainfeld, J., 253, 267, 268
Hajduk, P. J., 225
Hajibagheri, M. A. N., 149, 157
Hall, C. E., 253
Hall, L. R., 284
Halvas, E., 344
Hamaguchi, M., 259
Hamakubo, T., 306
Hamilton, A. D., 101, 104
Hamilton, E. H., 169, 170
Hamilton, J., 379
Hammersley, A. P., 350, 351, 352
Hamsten, A., 272
Hamstra, D., 410, 411, 412
Hance, J. E., 227
Handford, P. A., 406, 408, 409, 410, 412, 418, 419, 420, 423, 425
Hanefeld, F., 173
Hanein, D., 215
Häner, M., 114, 115, 124, 132
Hanna, L. S., 264
Hanna, R. B., 419, 423
Hansen, U., 345
Hansma, P. K., 356
Hanspal, M., 230, 231
Hanssen, E., 415, 440
Hantgan, R. R., 250, 252, 255, 261, 265, 268
Hanumanthaiah, R., 283

Harada, N., 429
Harbury, P. B., 19, 20, 43, 44, 55, 56, 60, 87, 89, 90, 91, 93, 96, 97
Hardcastle, R. A., 346
Harden, J. L., 99
Hardin, C. L., 272
Harding, C. R., 169
Harding, H. W., 170
Hariharan, I. K., 226
Harlan, J. E., 225
Harlos, K., 408, 409, 412, 418, 420, 423, 425
Haroun, R. I., 99
Harpel, P. C., 276, 277
Harper, P. S., 228
Harper, S., 218
Harrington, W. F., 303, 314
Harris, J. B., 219, 228
Harris, J. E., 260, 261
Harris, N. L., 70, 226
Harris, P., 230
Harris, P. C., 162
Harrison, P., 262
Harrison, S. C., 57
Harsh, M., 319
Hascall, V. C., 450
Haschemeyer, R. H., 253, 267
Hasegawa, K., 450
Haselgrove, J. C., 268
Haser, W. G., 284
Hasham, S. N., 427
Hashimoto, M., 346
Hashimoto, T., 19, 158
Hashimoto, Y., 170
Hashizume, H., 353
Haslam, R. J., 224
Hassinger, L., 179
Hassoun, H., 229, 230, 231
Haston, J. L., 359, 406, 411, 414, 417, 419, 420
Hatano, S., 222
Hatch, V., 60, 217
Hatsell, S. J., 152, 161
Hatton, Z., 229
Hatzfeld, M., 114, 117, 151, 159, 160, 161
Hatzinikolas, G., 416
Haudenschild, C. C., 274, 440
Hauptmann, G., 330
Hauptmann, R., 149, 154, 156, 159
Hauser, M. A., 219
Hausmanowa-Petrusewicz, I., 228
Haussinger, D., 377
Haverkate, F., 280
Havranek, J. J., 89
Hawiger, J., 279
Hawley-Nelson, P., 152
Hay, B. A., 226
Hay, S., 416
Hayashi, M., 450
Hayashi, T., 321, 325, 344, 358
Haydock, P. V., 168
Haynes, S. R., 117
He, L., 442
He, R., 427
He, T., 177
Heagerty, A. M., 416
Heath, R., 230
Hecht, M. H., 81
Hedbom, E., 345, 363
Hedlund, H., 357, 363
Hedman, K., 415
Heesemann, J., 45, 54
Heidemann, E., 302
Heim, R. A., 442
Heindrich, K., 228
Heinegard, D., 357, 358, 360, 363, 382, 386, 416
Heiska, L., 221
Helfand, B. T., 178, 179
Helfman, D. M., 59
Heljasvaara, T., 304
Heller, D. N., 256
Hellinga, H. W., 80, 81
Helliwell, T. R., 229
Hellner, K. A., 227, 228
Helseth, D. L., 347
Helseth, D. L., Jr., 347
Hemken, P. M., 180
Hemmings, B. A., 224
Hemmings, L., 214, 216, 222
Hempel, J., 344
Hempstead, P. D., 126
Henderson, J. M., 232
Henderson, M., 415, 422
Henikoff, S., 226
Henkel, W., 343, 349
Hennekes, H., 136
Henrich, S., 379
Henschen, A. H., 250, 253, 254, 255, 258, 267, 270, 282, 283
Henschen-Edman, A. H., 61, 282

Her, L. S., 180
Herbage, B., 344
Herbage, D., 344
Herder, G. V., 377
Hergt, M., 177
Hermans, J., 45, 265, 268
Hernandez, I., 263
Herrling, J., 114
Herrmann, H., 114, 115, 116, 120, 123, 124, 125, 131, 132, 135, 144, 154, 155, 156, 160, 167
Herrmann, R., 213
Hervio, L., 277
Hesketh, R., 415
Hesse, M., 116, 123, 166, 177
Hessel, B., 260, 267, 269, 272, 273
Heuser, J. E., 417, 419
Hew, F. H., 415
Hicks, L., 103
Hicks, M. R., 48, 55, 97
Higashino, K., 306, 327
Higazi, A. A.-R., 276
Higgins, D. L., 250
Higgins, M. K., 67
Highberger, J. H., 355
Hijikata, T., 156
Hikita, Y., 448
Hill, A. J., 174
Hill, C. P., 30
Hill, D. F., 213
Hill, R. B., 81
Hilliker, E., 256
Hilpela, P., 173
Hilton-Jones, D., 166
Hinds, M. T., 451, 452
Hindson, V. J., 426
Hinek, A., 416, 440, 443, 444, 448
Hinkle, R. T., 219
Hirai, H., 163
Hirako, Y., 154
Hirano, E., 416
Hirata, H., 280
Hirose, M., 304
Hirose, S., 261
Hirsch, M., 280
Hirschi, K. K., 448
Hitchcock-DeGregori, S. E., 25, 61
Hitomi, J., 353
Hiwada, K., 171
Hiyama, Y., 307, 318
Ho, C. L., 179
Hochmuth, R. M., 218
Hodge, A. J., 346, 347
Hodgens, K. J., 363
Hodges, J. A., 322
Hodges, R. S., 58, 59, 81, 84, 89, 90, 92, 93, 95, 97, 98, 101, 102, 103
Hodgkinson, J. L., 222
Hodson, N., 387
Hoenger, A., 219
Hoepner, B. L., 151
Hoeve, C. A., 449
Hoffman, A. S., 451, 452
Hoffman, E. P., 212, 228, 229, 376
Hoffman, S., 284
Hofmann, H., 323, 346
Hofmann, I., 120, 160
Hogan, K. A., 267, 278
Hogg, D., 267
Hohenester, E., 386, 387
Hoiczyk, E., 45, 54
Holberton, D. V., 19, 45, 53, 55
Holbrook, K. A., 169, 170
Holland, L. J., 260
Hollenbeck, J. J., 82, 101, 104
Hollosi, M., 19
Holm, L., 204, 209
Holman, K., 427
Holmback, K., 281
Holmes, D. F., 351, 411, 417, 418, 419, 420, 422, 425
Holmes, K. C., 60
Holmgren, S. K., 322
Holmskov, U., 416
Holstein, T. W., 377
Holton, J., 58
Holyst, T., 264
Holzbaur, E. L., 179
Homandberg, G. A., 267
Homes, D. F., 351, 354
Honjo, T., 416, 440
Hook, M., 306, 317, 319, 326
Hopkinson, S. B., 147, 152, 153, 156, 163
Horne, R. W., 447
Hornebeck, W., 448
Horrigan, S. K., 416
Horwitz, A. R., 211
Hoshino, T., 363
Hoskins, R. A., 226
Hosser, H., 231

Hotchin, N. A., 211
Ho-Tin-Noe, B., 277
Hou, G., 387
House-Pompeo, K., 326
Houser, S., 448
Houseweart, M. K., 144
Houtaud, A., 379
Howman, E. V., 165, 166
Hsiao, H., 444
Hu, C. H., 261
Hu, J. C., 88
Hu, W., 406, 408, 439
Hu, X., 284
Hu, Y. P., 222
Huang, C. C., 331
Huang, L., 451
Huang, X., 210, 224, 226
Huber, R., 379
Huber, S., 346
Hubner, C., 227, 228
Hudry-Clergeon, G., 265
Hudson, B. G., 330, 379
Hudson, D. L., 154
Hudson, L. G., 147, 160
Hudson, T. Y., 152
Huebner, N., 173
Huen, A. C., 147, 152, 160
Hugel, T., 449
Hughes, A. E., 152
Hughes, C., 67
Hughes, G. N. F., 361
Hughes, J. L., 439
Hughes, W. L., 250
Hughson, F. M., 99
Hugon, G., 213
Huiatt, T. W., 114, 136, 166, 180, 181
Hukins, D. W. L., 351
Hulmes, D. J., 317, 319
Hulmes, D. J. S., 13, 307, 344, 345, 346, 348, 351, 353, 356, 358, 363
Hulo, N., 13
Humphries, M. J., 382, 386, 387, 388, 407, 410, 412
Hunt, S., 388
Hunziker, T., 171
Husain-Chishti, A., 207
Huso, D. L., 428
Hutcheson, A. M., 174
Hutchinson, E. G., 100
Hutchison, C., 135, 136
Hutton, E., 150, 151
Huvenne, J. P., 447
Hwang, W. S., 361
Hyam, J., 151, 161
Hyde, T., 311
Hyland, J. C., 427
Hynes, R. O., 163, 226
Hyvonen, M., 224, 225

I

Ichikawa, I., 231
Idler, W. W., 117, 127, 132, 157, 158, 168, 170
Ievolella, C., 221
Igoe, F., 344, 358
Ihara, M., 429
Iizuka, H., 170
Ikeda, Y., 416, 440
Ikehara, Y., 261
Ikura, M., 223
Illidge, C., 386, 388
Ilsley, J. L., 22, 210, 224, 226, 227
Imada, K., 63
Imamura, M., 156, 165, 166, 222, 227
Imamura, Y., 344
Imanaka-Yoshida, K., 164, 213
Imanishi, T., 306, 327
Inada, H., 174, 175
Inagaki, M., 174, 175, 222
Indik, Z., 441, 442, 448
Inestrosa, N. C., 306, 326, 327, 328
Ingber, D., 163
Ingham, K., 254, 272, 274, 276
Inokuchi, Y., 19
Inouye, K., 319
Iolascon, A., 229
Iozzo, R. V., 30, 357, 358
Ip, W., 134, 150, 151, 166, 172, 180
Ireton, G. C., 81
Irvin, R. T., 95, 101
Irving, T. C., 350, 352
Irwin, M. I., 346
Ishibashi, R., 416
Ishida-Yamamoto, A., 158, 170
Ishidoh, K., 169
Ishihara, M., 172
Ishikawa, H., 156
Ishikawa-Sakurai, M., 227

Ishimaru, S., 448, 451
Islam, N., 363
Isogai, Z., 416, 440
Itakura, H., 306, 327
Itano, T., 416
Ito, H., 324
Ito, M., 172
Ito, S., 448, 451
Itoh, T., 346
Ivaninskii, S., 99, 123
Iwamatsu, A., 159
Iwanaga, Y., 416, 440
Izaguirre, G., 222
Izawa, I., 174, 175
Izmiryan, A., 166
Izumi, Y., 365

J

Jaalinoja, J., 357
Jacenko, O., 387
Jackson, D. G., 282
Jacob, M.-P., 439, 445, 447, 448, 451
Jacobsen, S., 179
Jacobsson, M., 306, 323
Jagadeeswaran, P., 283
James, M., 224, 226, 227
Janda, L., 155, 159
Jandrot-Perrus, M., 262
Jang, S.-I., 19, 59, 114, 124, 131, 132, 134, 181
Janmey, P. A., 179, 216, 273
Jarchau, T., 19, 20, 21, 22, 124
Jarolim, P., 229, 230
Jarrett, H. W., 222
Jaunin, F., 150, 151, 152, 154
Jeffersohn, E. A., 302
Jefferson, J. J., 146, 147, 151, 153
Jelinski, L. W., 307, 316, 318
Jenkins, C. L., 302, 315, 321, 322
Jenny, M., 21, 54, 124, 309
Jensen, D. E., 443
Jensen, K. D., 161
Jensen, N. J., 281
Jensen, S. A., 414, 445, 447
Jepsen, K. J., 358
Jermiin, L. S., 441
Jesoir, J. C., 350
Jewell, D. P., 174
Jimenez, S., 319
Jimenez, S. A., 324, 418
Jin, L. H., 361
Jirouskova, M., 279
Joachimiak, A., 210, 224, 226
John, K. M., 229, 230, 231
Johnson, C. A., 230, 231
Johnson, L. D., 117, 132
Johnson, M. A., 219, 228
Johnson, P. F., 61, 88
Jondeau, G., 429
Jones, E. Y., 347, 348
Jones, F. S., 284
Jones, J. C., 147, 152, 153, 156, 163
Jones, J. C. R., 154
Jones, L. N., 127, 132
Jones, P., 204
Jones, S. J., 226
Jonkman, M. F., 147
Jorcano, J. L., 176
Jörgl, A., 156
Joseph, A., 85, 97, 98, 100, 103
Joseph, C., 224
Jovanovic, J., 408, 409, 412, 418, 420, 423, 425
Judge, D. P., 428
Julien, J. P., 153, 179
Jun, G., 25, 61, 255, 258, 270
Jung, C., 179
Jung, D., 227, 229
Junien, C., 427, 429
Junius, F. K., 57

K

Kaariainen, H., 228
Kadler, K. E., 326, 327, 328, 351, 354, 358, 382, 386, 387, 388, 417, 418, 419, 420, 422, 423
Kagan, H. M., 439, 445
Kagawa, H., 25
Kagawa, M., 379
Kahana, E., 221
Kahari, V. M., 440, 442, 443
Kahl, S. D., 212, 213
Kahsai, T. Z., 379
Kaikoku, T., 181
Kaitila, I., 427
Kajava, A. V., 29, 30, 90, 96, 98

Kajii, T., 429
Kajikawa, T., 355
Kajitani, M., 451, 452
Kakudo, M., 303, 304, 329
Kalikhevich, V. N., 265
Kalimo, H., 175
Kalinin, A. E., 157, 158
Kalluri, R., 379
Kam, E., 168
Kamawal, A., 57
Kaminski, M., 272, 274
Kamiyama, R., 355
Kammerer, R. A., 21, 54, 59, 84, 100, 124, 126, 309
Kane, D., 359
Kannus, P., 361, 365
Kant, J. A., 260
Kanzaki, A., 230
Kaplan, C., 262, 281
Kaplan, J.-C., 228
Kaplan, J. M., 231, 232
Kaplan, L. S., 282
Kapprell, H.-P., 160
Kaptein, R., 54
Karabinos, A., 114, 134
Karashima, T., 157, 158
Karn, J., 25, 45, 51
Karpe, F., 272
Kartenbeck, J., 149
Kartha, G., 18, 302, 307, 312
Kassner, A., 323, 345
Katsumata, O., 344
Katz, S. I., 152
Kaufmann, S. H., 174
Kaulfus, P., 172
Kawabe, Y., 306, 327
Kawashima, H., 304
Kawashima, S., 169
Kay, B. K., 226
Kay, C. M., 58, 84, 90, 102, 103
Kazerounian, S., 157, 158
Kazmierski, S. T., 172, 173
Kazuya, H., 324
Keane, P. M., 256, 282
Keating, A. E., 58, 89
Keating, L., 168
Keating, M. T., 438, 448
Keeffe, E. B., 174
Keeley, F. W., 439, 444, 447, 451, 453
Keene, D. R., 306, 382, 406, 408, 411, 415, 416, 417, 428, 439, 440
Keep, N. H., 214, 215, 216, 217
Kehl, M., 283
Keijer, J., 276
Kellermann, J., 53, 60
Kelly, A. M., 213
Kelly, G., 224
Kelly, T. R., 100
Kelly, V., 411, 412, 413
Kelsell, D. P., 152, 161
Kendrick-Jones, J., 207, 210, 212, 213, 214, 215, 216, 217, 219, 222, 227
Kennan, A. J., 89
Kennedy, C. R., 232
Kennedy, S. P., 209, 220, 231
Kent, H., 60
Kere, J., 221
Kerr, T. P., 228
Keski-Oja, J., 415, 416
Kessler, E., 363
Khaleduzzaman, M., 13, 17, 159
Khalil, P. N., 451
Khandekar, M. J., 147, 160
Khun, K., 379, 380
Khuon, S., 180
Kido, N., 154
Kieff, E., 181
Kiehart, D. P., 206, 208
Kielty, C. M., 303, 304, 381, 382, 386, 387, 388, 406, 407, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 422, 423, 425, 426, 428, 439, 440, 444
Kielty, K., 382
Kihara, S., 304
Kikuno, R., 406
Kilchherr, E., 323
Kilpatrick, K. E., 349
Kilpelainen, I., 357
Kilroy, S. M., 451
Kim, E., 59, 161, 162
Kim, I.-G., 169, 170
Kim, K. H., 25, 61
Kim, P. S., 19, 20, 43, 44, 47, 55, 56, 60, 61, 82, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 99, 101, 104
Kim, S. H., 231, 232
Kimball, J. R., 168
Kimura, J. H., 358
Kimura, K., 231

Kimura, T., 304
Kindt, J. H., 356
King, D. S., 95, 96
King, E., 389
King, G. F., 57, 349
King, W., 228
Kingston, H. M., 228
Kinoshita, A., 429
Kirkpatrick, A., 320, 322, 324, 325, 326, 327, 328, 331
Kirschner, M. W., 125
Kishida, Y., 319
Kishino, T., 429
Kishore, U., 304, 306
Kitagawa, M., 221
Kitahama, S., 416
Kitamura, T., 355
Kiuru, A., 228
Kivirikko, K. I., 303, 304, 329
Kiyokawa, C., 158
Kiyonaga, S., 363
Kiyono, T., 175
Klasen, C., 156
Klauser, S., 56, 88
Klaushofer, K., 351
Klein, C. J., 228
Klein, T. E., 331
Klement, J. F., 158
Klemm, J. D., 61, 85, 87, 88, 90, 91, 94
Klug, A., 29
Klymkowsky, M. W., 136
Knight, A. E., 210, 213, 227
Knight, B., 211
Knight, C. G., 308, 309, 312, 318, 326, 327, 328, 329
Knight, D. P., 351, 388, 389
Knorzer, E., 353, 366
Knott, V., 408, 409, 410, 412, 418, 420, 423, 425
Knowlton, J. R., 30
Knowlton, R., 153
Knupp, C., 377, 382, 383, 385, 388, 389, 390, 391, 392, 393
Kobatake, E., 452
Kobayashi, M., 346, 363
Kobayashi, R., 416
Kobayashi, Y., 319, 322
Kobe, B., 29, 30
Kobuke, K., 416, 440
Koch, E. A., 114
Koch, M., 344
Koch, M. H., 353
Koch, M. H. J., 353
Kocher, O., 162
Kodama, T., 306, 327
Koehn, J. A., 280
Koenig, M., 204, 212, 228
Kohl, A., 59
Kohn, W. D., 81, 93
Kohno, K., 159
Koide, H. B., 224
Koide, T., 327
Kojic, S., 221
Kojima, S., 98
Kollman, J. M., 61, 255, 258
Kollman, P. A., 331
Kolmerer, B., 172, 173
Kombrinck, K. W., 282
Kominami, E., 169
Koonin, E. V., 80
Koopman, J., 280
Kopan, R., 181
Kopecek, J., 99
Kopp, D. S., 151, 160
Korenbaum, E., 214, 222
Korge, B. P., 117, 118, 133
Korkko, J., 326, 327
Kornacker, I., 156
Kornalik, F., 260
Koronakis, E., 67
Koronakis, V., 67
Korsholm, B., 275
Korte, C. C., 206, 208
Kos, C. H., 232
Koschinsky, M. L., 276, 277
Kost'an, J., 155
Kostelansky, M. S., 258, 259, 269
Koster, A., 417, 420
Koster, A. J., 351, 354, 417, 418, 419, 420, 422, 423
Koster, J., 147, 152, 153, 155
Kostka, G., 416
Kostrewa, D., 100
Koteliansky, V., 163
Kothary, R., 152, 153
Kotovych, G., 347, 349
Kotula, L., 19, 204
Kouklis, P., 150, 151, 152, 153
Kowalczyk, A. P., 147, 149, 151, 160, 178
Kowalczyk, C., 55

Koyama, T., 284
Kozel, B. A., 444
Kramer, D., 155
Kramer, L., 30
Kramer, R. Z., 308, 309, 312, 314, 315, 316, 318
Krams, S. M., 174
Kranewitter, W., 214
Kraulis, P. J., 85
Krepler, R., 154, 156
Kretsinger, R. H., 222, 223
Krieg, T., 163
Krimm, S., 312
Krishna, S. S., 29
Krissansen, G. W., 19
Kristjansson, G. I., 160
Kriz, W., 231, 232
Kruger, M., 172
Krylov, D., 57, 89
Ku, N. O., 124, 174, 175, 176, 177
Kubitscheck, U., 415
Kucera, J., 265
Kuchel, P. W., 439, 445
Kucich, U., 416, 448
Kudoh, J., 19
Kudryk, B., 261
Kuechle, M. K., 169
Kuehl, P. M., 226
Kuhlman, B., 81
Kuhlman, P. A., 216
Kuhlmey, A., 229
Kuhlmorgen, B., 53, 60
Kuhn, F., 171
Kuhn, K., 18, 323, 327, 346, 377, 379
Kuhn, S., 177
Kühnel, K., 19, 20, 21, 22, 124
Kuikman, I., 155
Kuliwaba, J. L., 416
Kumar, K., 89
Kumar, R., 274
Kumaratilake, J. S., 414, 415
Kumasaka, T., 63
Kummerfeld, S. K., 441
Kunimatsu, M., 13, 17, 159
Kunkel, L. M., 165, 166, 204, 212, 227, 228
Kuo, H.-J., 382, 417
Kuo, W. L., 221
Kurihara, H., 232
Kurihara, Y., 306, 327
Kuriki, Y., 98
Kuriowa, K., 312
Kurochkina, N., 276
Kvansakul, M., 386, 387
Kwak, J., 302
Kwan, A. P., 387
Kwan, A. P. L., 388
Kwok, S. C., 97
Kyogoku, Y., 319

L

Labeit, S., 172, 173
Lack, J., 410
Lacombe, P. J., 223
Lacour, J. P., 163
Lahav, M., 229
Lai, T.-S., 250, 270, 271, 272, 282
Laine, R. A., 256
Laing, J. H., 354
Laing, N. G., 173
Laio, G., 178
Laki, K., 250
LaMantia, C., 358
Lamb, J. A., 150, 151
Lanar, D. E., 25
Landis, W. J., 363
Landon, F., 174
Landschulz, W. H., 61, 88
Landwehr, R., 21, 54, 59, 84, 124, 309
Lane, B., 135, 136
Lane, B. P., 273
Lane, E. B., 154, 155, 174, 176
Lanfranchi, G., 221
Langa, F., 163
Langbein, L., 170, 171
Langendijk-Genevaux, P. S., 13
Langer, B. G., 256
Langer, R., 101, 448, 452
Langeveld, J. P., 379
Langford, G. M., 179
Langley, R., 19
Langley, R. C., Jr., 167
Langowski, J., 125
Langridge, R., 265
Languino, L. R., 278
Lanina, N. F., 98
Lankinen, H., 357
Lanzi, G., 153
Laplaud, M., 277

Lariviere, R. C., 179
Lass, J. H., 358
Lassing, C., 171
Lassmann, H., 156
Lau, K. S., 166
Laudano, A. P., 279
Lauer-Fields, J. L., 319
Laukaitis, C., 211
Laurila, P., 19, 204
Lausen, M., 416
Laveder, P., 221
Lavigne, P., 58, 59, 84, 89, 90, 103
Lavoie, T. B., 95
Law, C. M., 151
Law, R., 218
Law, S. F., 163
Lawler, J., 229
Lawrence, P. A., 343
Lawrence, S. O., 259
Lawrie, A. S., 262
Laws, W. R., 318
Lawson, D., 155, 156
Lazarides, E., 166, 168, 212, 230
Lazebnik, Y. A., 174
Lazzarini, R. A., 19
Le, T. C., 232
Leapman, R. D., 96, 98
Leavesley, D. I., 415
Leavitt, J., 172
Lebart, M. C., 216
Le Bourdelles, S., 427
Lechner, J. H., 347
Leclerc, N., 153
Ledo, I., 443
Lee, D. L., 59, 84
Lee, E. C., 284
Lee, H. Y., 222
Lee, J., 452
Lee, J. C., 174, 218
Lee, J. H., 59, 124, 131
Lee, M.-K., 114, 181
Lee, S.-C., 169, 170
Lee, S. S., 408, 409, 412, 418, 420, 423, 425
Lee, T. C., 450
Lee, V. M., 179
Lee, V. M.-Y., 19
Lee, Y. C., 256
Lee, Y. M., 174
Lefebvre, F., 448, 451, 453
Leger, J. J., 213
Legrand, P., 447
Lehman, W., 215, 217
Lehner, M., 174
Lehto, V. P., 204, 209, 223
Leichtfried, F. E., 154, 155
Leigh, I. M., 152, 154, 155, 161
Leikin, S., 315, 316
Lemaitre, B., 226
Lemieux, L. I., 232
Lemmon, M. A., 224, 225
Lendahl, U., 128
Lenk, U., 227, 228
Le Rumeur, E., 221
LeSaux, V., 13
Leslie, J. G., 438
Leslie, N. J., 356
Lesty, C., 275, 277
Leterrier, J. F., 179
Lethias, C., 306
Leto, T., 19
Leto, T. L., 204
Le Trong, I., 270
Leube, R. E., 151
Leung, C. L., 146, 147, 151, 152, 153, 155
Leung, D. Y. M., 448
Leung, E., 19
Leung, L. L., 274
Levi-Minzi, S. A., 440
Levine, B. A., 216
Levitt, M., 40, 45, 83
Lewis, M. B., 330
Lewis, S., 226
Lewis, S. D., 264
Lewis, S. P., 169
Li, B., 172, 361
Li, D., 218
Li, D. Y., 438, 448
Li, J., 226
Li, K., 149
Li, M., 312, 318, 322
Li, P. W., 226
Li, T., 439, 445
Li, X., 162, 220
Li, Y., 61, 172
Li, Y. T., 256
Li, Z., 166, 174, 226
Li, Z. Y., 416
Liao, C. X., 136

Liao, J., 174, 176
Liddington, R. C., 308, 309, 312, 318, 328
Lidov, H. G., 165
Liebowitz, D., 181
Liechti-Gallati, S., 228
Liem, R. K., 114, 134, 146, 147, 151, 152, 153, 155, 174, 179
Lieska, N., 154
Light, N. D., 356
Lill, H., 264
Lillie, M. A., 446, 453
Lim, W. A., 226
Lin, C. S., 172
Lin, G., 406, 411, 413
Linders, M., 276
Lindlof, M., 228
Linhardt, R. J., 452
Linn, H., 224, 226, 227
Linnenbach, A., 19, 204
Linsenmayer, T. F., 344, 358
Lipas, A. N., 82, 86
Lishko, V. K., 279
Litjens, S. H., 155
Litowski, J. R., 95, 101
Little, C. D., 440
Littleton, J. T., 226
Litvinov, S. V., 162
Litvinovich, S. V., 254, 255, 258, 265, 274
Liu, C. G., 154
Liu, C. Y., 280
Liu, H., 282
Liu, J., 219
Liu, J. M., 443
Liu, J. N., 277
Liu, N., 56
Liu, N. K., 88
Liu, S. C., 229
Liu, W. G., 416
Liu, X., 180, 439, 445
Liu, X. H., 349
Liu, X. S., 80
Liu, Y. C., 283
Ljungquist, A., 416, 439
Lloyd, S. J., 63
Locardi, E., 302
Lock, J. B., 269
Loewy, A. G., 270
Loeys, B., 427
Lofberg, H., 416, 439
Lofvenberg, L., 223
Logan, J., 96
Loh, N. Y., 165, 213
Lomas, A., 407, 410, 412
Lombardi, A., 81, 82, 93
Long, C. G., 330
Long, M. M., 445
Lonsdale-Eccles, J. D., 168, 169
Looger, L. L., 80, 81
Loomis, P., 179
Loos, M., 330
Lorand, L., 268, 270, 271, 273
Loranger, A., 173, 175, 177
Lord, S. T., 258, 259, 264, 267, 269, 270, 278, 279, 280
Loscalzo, J., 277
Lottspeich, F., 283
Louache, F., 262
Loukinov, D. I., 279
Lounes, K. C., 259, 278
Louvard, D., 177
Love, D. R., 213, 219, 228
Love, R., 178
Lovejoy, B., 58, 93, 94, 101
Lovely, R. S., 264
Lowthert, L. A., 174
Lozano, P. R., 416, 440
Lu, J., 304, 306
Lu, R., 13, 17, 159
Luber, K., 149, 154
Lubin, B., 230
Lucas, F., 15
Lucas, J., 283
Luciano, L., 382
Lucic, D., 349
Luckman, S. P., 387
Ludescher, R. D., 318
Ludosky, M. A., 161
Ludwig, T., 323, 345
Lue, R. A., 206, 210
Lukinova, N. I., 254, 265
Lukomski, S., 306
Lumb, K. J., 56, 60, 84, 89, 90, 92, 93, 95, 105
Lund, L. R., 282
Lundberg, S., 223
Luo, Z. J., 177
Luong, T.-T., 388
Lupas, A., 14, 45, 47, 49, 51, 53, 54, 55, 60, 82, 86, 87, 90, 91
Lupski, J. R., 229

Luse, S. A., 382
Lustig, A., 59, 84, 99, 123, 132
Luther, P. K., 219, 382, 383, 389
Luthert, P. J., 382
Lützen, O., 275
Luu, H. H., 150, 151
Lux, S. E., 229, 230, 231
Ly, B., 259
Lynch, J. R., 406
Lynch, N., 416

M

Ma, L., 168, 180
Ma, R. I., 416
Maatta, A., 157, 158
Mabuchi, K., 172
MacBeath, J. R., 358
MacDonald, R. I., 218
Macias, M. J., 224, 225
Maciver, S. K., 214, 222
Mack, J. W., 117, 127, 169
MacKenzie, R., 177
MacNulty, E., 150, 155
Macosko, C. W., 452
MacPhee, C. E., 81, 98, 99
MacRae, T. P., 14, 18, 21, 23, 44, 58, 117, 124, 126, 127, 303, 307, 308, 314, 316, 319, 321, 322, 350, 351, 353
Maddox, B. K., 417
Madrazo, J., 255, 258, 263
Maercker, C., 154
Magatti, D., 269
Magee, A. I., 151, 161
Magin, T. M., 116, 123, 166, 174, 177
Magnuson, T., 358
Mahoney, M. G., 158
Maimone, M. M., 166
Main, E. R., 59
Majsterek, I., 324
Majumdar, I., 29
Makogonenko, E., 274, 276
Makowski, L., 272
Maksymiw, R., 221
Malacrida, S., 152
Malashkevich, V. N., 89, 92
Malone, J. P., 348
Manabe, M., 169, 170
Mangeat, P. H., 167
Maniglia, J. V., 451, 452
Maniotis, A., 163
Mann, K., 379
Mann, K. G., 269
Mannucci, P. M., 280
Mannweiler, K., 155
Mant, C. T., 59, 89, 97
Mantero, S., 360, 366
Maraschio, P., 153
Marceau, N., 173, 175, 177
Marchesi, S. L., 230
Marchesi, V. T., 19, 167, 204, 206
Marchini, M., 346, 349, 353
Marchisio, P. C., 278
Marcovina, S. M., 276
Marder, V. J., 250, 252, 255, 261, 262, 264, 275
Mardon, H. J., 408, 409, 412, 418, 420, 423, 425
Marechal, J., 230
Marekov, L. N., 19, 23, 59, 114, 115, 117, 120, 121, 122, 123, 124, 125, 127, 128, 129, 131, 132, 134, 135, 157, 158, 168, 169, 170, 181
Maret, G., 265
Marigo, V., 440
Marini, J. C., 330
Marini, R., 448
Markely, J. L., 322
Markle, D. R., 218
Marko, M., 363
Markova, N. G., 17, 168
Marsden, R. F., 213
Marshall, J., 19, 45, 53
Marson, A., 407, 414, 415, 422, 444
Martin, B. D., 452
Martin, J., 72
Martin, P., 163
Martin, R., 363
Martin, S. L., 453
Martinelli, R. A., 264
Martinez, J., 213, 279
Martino, M., 451
Martin-Rendon, E., 165, 180
Maruyama, K., 172
Masaki, T., 211
Mason, T. L., 451, 452
Masse, J. M., 262
Masters, E. I., 30
Masuda, T., 355

Mathews, M. B., 448
Mathieu, M., 152, 153
Mathis, B. J., 231, 232
Matinez, J., 256, 282
Matoltsy, A. G., 149
Matsuchima, N., 365
Matsuda, M., 279, 280, 304
Matsudaira, P., 172, 214, 215
Matsueda, G. R., 95
Matsuka, Y. V., 254, 272
Matsumoto, A., 306, 327
Matsumoto, N., 429
Matsumura, K., 213
Matsumura, T., 163
Matsuo, N., 13, 17, 159
Matsuzawa, Y., 304
Mattei, M.-G., 153, 379, 440
Mattioli, F., 448
Matyas, G., 427
Maurer, P., 357
Mauviel, A., 443
Maxwell, A. L., 445
Mayer, B. J., 224, 226
Mayer, J. E., Jr., 451
Mayne, R., 346
Maynell, L. A., 136
Mayville, P., 308, 309, 312
Mazzarella, L., 309, 314, 322
McAdams, E., 324
McAlinden, A., 307
McAndrew, P., 228
McBride, D. J., 343, 346
McBride, O. W., 17, 117, 118, 133, 168, 260
McCauley, P., 221
McClain, D. L., 104, 105
McCrea, P. D., 159
McDonagh, J., 250, 253, 270, 272, 274, 279
McDonald, M., 19
McElhinny, A. S., 172, 173
McEwen, B. F., 363
McGavin, S., 302, 307
McGee, C., 158
McGough, A., 215, 216
McGrath, J. A., 154
McGrath, K. P., 99
McIntosh, S., 230
McKay, D. B., 61
McKay, R. D. G., 128
McKenna, K. E., 152
McKeon, F. D., 125
McKern, N. M., 127
McKnight, S. L., 61, 88
McLachlan, A. D., 14, 24, 25, 26, 45, 51, 57, 90, 126
McLean, W. H. I., 154, 156
McLeod, D., 406
McMillan, C., 218, 230
McMillan, J. R., 147, 154
McMillan, R. A., 451
McMillan, R. P. A., 451
McPherson, D. T., 451, 452
McPherson, J. D., 416
McPhie, P., 157, 158
McRorie, D. K., 58, 93, 94, 101
Meade, T. W., 272
Meadows, R. P., 225
Mecham, L., 415
Mecham, R. P., 406, 408, 410, 412, 413, 414, 415, 416, 417, 419, 438, 439, 440, 443, 444, 445, 446, 448, 450
Mechling, D. E., 411
Medved, L. V., 254, 255, 258, 263, 267, 272, 274, 276, 277, 279, 283
Meek, K. M., 346, 360
Meh, D. A., 250, 264, 270
Mehrani, T., 122, 128, 129
Mei, G., 169, 170
Meier, B. H., 100
Meier, M., 84, 99
Meier, S., 377
Meiner, V., 152
Meiyappan, M., 379
Mejean, C., 216
Melanson, L., 254, 257
Melin, M., 250
Melino, G., 169, 170
Mellody, K., 407, 414, 415, 422, 444
Melnik, T. N., 98
Menache, D., 262
Mendell, J. R., 228
Mendler, M., 346
Meneguzzi, G., 163
Meng, G., 228
Meng, J.-J., 150, 151
Mengarelli-Widholm, S., 357, 363
Menninger, J. C., 230
Mercer, R., 415
Merckling, A., 163
Merkulov, S. M., 279
Merle, C., 317, 319

Mertens, C., 160
Metzger, S., 152
Meuli, M., 415
Meyer, D. E., 451
Meyer, J., 227
Michalak, M., 227
Michaud, J. L., 232
Michel, J. B., 451
Michie, S. A., 174, 176
Michna, H., 360
Michnick, S. W., 58, 104
Michon, A. M., 215
Micklatcher, C., 82
Middlesworth, W., 228
Middleton-Price, H. R., 228
Midura, R. J., 450
Mies, B., 166
Migchelsen, C., 314
Migliorini, M. M., 272
Mihalyi, E., 260, 264
Mikami, A., 178
Mikhailenko, I., 57
Mikhailov, A., 175
Miki, K., 19
Mikkola, H., 270
Milbradt, A. G., 377
Milewicz, D. M., 410, 411, 412, 413, 427, 450
Milla, M., 47
Miller, A., 314, 316, 346, 347, 348, 349, 350, 351, 352, 353, 356
Miller, R., 225
Milner, R. E., 227
Mimmrich, V., 160
Mimura, N., 216
Minamisawa, S., 416, 440
Minetti, C., 213
Minin, A., 178
Minor, D. L., Jr., 60
Minosse, C., 376
Minuth, W. W., 451
Mir, B., 150, 155
Miraglia, C. C., 270
Miraglia del Giudice, E., 229
Miravet, S., 160
Mirshahi, M. C., 280
Mischke, D., 117, 118, 133
Misof, K., 366
Misra, S., 226
Mital, R., 214
Mithieux, S. M., 453
Mitkevich, O. V., 264
Mittl, P. R. E., 56, 88
Mittnacht, S., 176
Miura, K., 98
Miyahara, M., 221
Miyasaka, M., 304
Miyata, T., 452
Miyazaki, T., 306, 327
Mizuguchi, T., 429
Mizuno, K., 321, 325, 344
Mizuno, Y., 165, 166
Mizushima, K., 166
Mizutani, H., 363
Mlodzik, M., 284
Moaddel, M., 264, 270
Mochalkin, I., 61
Mock, M., 306
Mockros, L. F., 268, 273
Moen, J. L., 270
Mohan, P. S., 179
Mohandas, N., 221
Moir, A. J., 216
Moir, R. D., 144, 174, 178, 180
Moitra, J., 89
Moll, J. R., 95
Mollenhauer, J., 349
Momoi, T., 175
Momota, R., 379
Monaco, A. P., 204, 212, 213, 228
Monaghan, P., 343
Mondragon, A., 218
Monera, O. D., 58, 102, 103
Mongiat, M., 440
Mongiu, A. K., 152
Monni, O., 416
Montalescot, G., 275, 277
Monteiro, M. J., 128
Montevecchi, F. M., 360, 366
Montezin, M., 169
Moon, R. T., 230
Mooney, S. D., 331
Moore, E., 415
Moores, C. A., 214, 215, 216, 217
Moorhead, G. B., 172, 176
Morasso, M. I., 59, 122, 124, 128, 129, 131
Morgan, A., 407, 410, 411, 412, 414, 415, 420, 422, 444
Morgan, F. J., 280
Morgan, G., 228

Morgan, J. G., 261
Morgelin, M., 323, 345, 357, 382, 386, 416
Mori, G., 416, 447, 448
Mori, O., 158
Moriizumi, K., 355
Morimoto, C., 163
Morisaki, H., 429
Morisaki, T., 429
Morita, K., 280
Morle, L., 230
Mornet, D., 213, 222
Morocutti, M., 353
Moroder, L., 328, 329, 377
Morris, C., 19
Morris, C. J., 355
Morris, E. P., 25, 389
Morris, G. E., 229
Morris, M. B., 439, 445
Morrison, D. K., 226
Morrow, J. S., 209, 220, 221, 231
Morse, D. E., 356
Morton, L. F., 326, 327, 329
Mosavi, L. K., 60
Moser, H., 228
Moses, M. A., 451
Mosesson, M. W., 250, 253, 260, 262, 263, 264, 267, 268, 270, 272
Mosler, E., 353, 366
Mostad, S., 169
Mountford, R. C., 229
Moyer, M., 228
Muchowski, P. J., 174
Mucignat, M. T., 440
Mücke, N., 125
Mueller, M., 171
Mui, S., 61
Muilenburg, A., 427
Muiznieks, L. D., 447
Mukai, H., 221
Mulford, D. J., 250
Mulholland, D., 387
Muller, C. R., 228
Müller, J. F., 21
Müller, K. M., 58
Müller, M. F., 267
Müller, P. K., 409, 412, 413, 416
Müller, S., 51, 53, 60, 114, 115
Muller, S. A., 100
Mullin, J. L., 264
Mundel, P., 231
Mundy, G. R., 415
Mungall, C., 226
Mungiguerra, G., 440
Munro, P. M., 382, 383
Muntoni, F., 173
Murakami, T., 156, 306
Murata, S., 156
Murata, T., 181, 412
Murdoch, A. D., 30
Murphy, G., 426, 428
Murphy, P., 228
Murray, J., 229, 230, 231
Musacchio, A., 224, 225
Muse, S. V., 206, 208
Muskett, F. W., 224
Myers, L., 428, 429
Mykkanen, O. M., 221
Myllyharju, J., 303, 304, 329
Myszka, D. G., 103

N

Nagase, H., 319
Nagase, T., 406
Nagashima, S., 63
Nagaswami, C., 256, 259, 272, 275, 276, 277, 278, 280
Nagata, K., 327
Nagel, R. L., 230
Nagelkerke, J. F., 275
Nair, C. H., 273
Naito, I., 159, 379
Nakagawa, H., 172
Nakajima, D., 406
Nakamae, K., 452
Nakamoto, T., 163
Nakamura, H., 306, 327, 448
Nakamura, T., 416, 440
Nakamura, Y., 429
Nakanishi, S., 355
Nakayama, M., 406
Nakayama, S., 223
Nakazawa, T., 322
Namba, K., 63
Namba, Y., 172
Napolitano, E. W., 134, 174
Nasseri, B. A., 451
Nath, N., 452
Nautiyal, S., 95

Nava, A., 152
Nawrotzki, R., 165
Neerman-Arbez, M., 281
Neilson, E. G., 330
Nelson, C. R., 226
Nelson, W. J., 159
Nemethy, G., 315, 348, 349
Nemetschek, T., 353, 366
Nemetschek-Gansler, H., 353, 366
Neptune, E. R., 429
Nesheim, M. E., 276
Ness, S. A., 162
Netzer, K. O., 379
Nevett, C., 415
Newbern, E. C., 206, 208
Newey, S. E., 165, 212, 213, 227
Ngo, K., 361
Nguyen, M. D., 179
Nham, S. U., 262
Ni, J., 19
Nichol, M. C., 166
Nicholl, I. D., 173
Nicholson, L. V., 219, 228
Nicol, A., 452
Nie, Z., 158
Niemeyer, C. M., 100
Nieminen, M. I., 177
Nietlispach, D., 224
Nieuwenhuizen, W., 254, 267, 275, 276
Nievers, M. G., 149, 152, 155
Nigg, E. A., 136
Niikawa, N., 429
Niklason, L. E., 448
Nikolic, B., 150, 155
Nilges, M., 48, 49, 57, 61, 224, 225
Nilles, L. A., 149, 150
Ninomiya, Y., 379, 387
Nirunsuksiri, W., 168, 169
Nishi, K., 158
NIshi, Y., 322
Nishida, W., 171
Nishikawa, T., 19, 156, 158
Nishiuchi, Y., 322
Nishiyama, Y., 181
Nishizawa, M., 174, 175
Nishizawa, Y., 154, 181
Nislow, C., 136
Nitsch, M., 53, 60
Niu, C. H., 307, 318
Nixon, R. A., 179
Njolstad, P. R., 229, 230
Nobiling, R., 232
Noble, M., 225
Noblesse, E., 439, 445
Nodder, S., 163
Noegel, A., 222
Noelken, M. E., 379
Nomura, K., 153
Nooren, I. M. A., 54
Norgett, E. E., 152, 161
North, A. C. T., 117, 127, 129, 133, 302, 307, 350
North, A. J., 151, 161, 172
North, K. N., 231, 232
Norwood, F. L., 214, 215, 216, 217
Nossel, H., 262
Notbohm, H., 323, 345, 409
Novak, M. R., 151
Novotny, J., 95
Nugent, M. A., 440, 443
Nuovo, G., 262, 281
Nusse, R., 159
Nuttall, A., 224, 226, 227
Nuytinck, L., 427

O

Oakley, M. G., 82, 92, 101, 103, 104, 105
Oas, T. G., 99
Oberbaumer, I., 377
O'Brien, K. F., 227
Oda, H., 448
Oda, K., 261
Odajima, S., 366
Odent, S., 173
Odermatt, E., 380, 382, 386
O'Donoghue, S. I., 49, 57
Odrljin, T. M., 259, 274
Oehler, S., 154, 156, 159
Oexle, K., 227, 228
O'Farrell, P. H., 226
Offer, G., 44, 48, 55
Ogawa, D., 304
Ogawa, H., 169
Ogawa, S., 163
Ogilvie, M. L., 274
O'Guin, W. M., 170
Ohanian, V., 210, 211, 223
Ohara, O., 153, 406

Ohashi, K., 172
Ohkubo, T., 322
Ohlendieck, K., 213
Ohnishi, T., 445, 447, 450
Ohta, T., 429
Ohtakara, K., 174
Ohtsuka, K., 174, 181
Okamoto, K., 166, 416
O'Keefe, E. J., 149, 169, 170
Oklu, R., 415
Okuaki, T., 450
Okumura, M., 153, 154
Okumura, N., 267, 279
Okuyama, K., 303, 304, 309, 329
Olausson, T., 219
Oldberg, A., 358, 360
O'Leary, J. M., 410
Olejniczak, E. F., 225
Olin, A. I., 416
Olsen, B. R., 345
Olsen, E. N., 416
Olson, E. N., 440
Omary, M. B., 144, 174, 175, 176, 177
O'Neil, K. T., 93
O'Neill, G. M., 163
Ono, R. N., 409, 410, 413, 416, 428, 440
Ono, Y., 221
Onoda, K., 452
Oohashi, T., 379
Ooyama, T., 448
Oppers-Walgreen, B., 213
O'Rear, J., 379
Orgel, J. P., 348, 349, 350, 352
Oriel, B., 254, 257
Orlova, A., 215, 217
Ornstein-Goldstein, N., 440, 441, 442
Orrison, B. M., 330
Ortega, J., 30
Ortolani, F., 346, 349
Ortonne, J. P., 163
Osakabe, T., 450
Oschkinat, H., 224, 225
O'Shea, E. K., 61, 85, 87, 88, 90, 91, 93, 94
Oshima, R. G., 174
Osmanagic-Myers, S., 157
Osornio-Vargas, A., 355
Otani, Y., 13, 17, 159
Otey, C. A., 212
Ottani, V., 353, 354, 360
Otter, A., 347, 349
Ottersbach, K., 224, 226, 227
Otto, J. J., 211
Otvos, L., 19
Ouchi, T., 416, 440
Ouyang, P., 161
Owaribe, K., 149, 153, 154, 160, 163
Owen, C., 363
Owens, D. W., 174
Ozawa, E., 165, 227
özbek, S., 21, 377

P

Pachter, J. S., 174
Padilla, J. E., 99
Paganini, E., 269
Page, D., 228
Page, I. H., 269
Pagni, M., 13
Paige, M. F., 355
Palascak, J. E., 256, 282
Palek, J., 229, 230, 231
Palka, H. L., 147, 151, 160
Pallari, H. M., 175
Pallavicini, A., 221
Palliser, C. C., 126
Palumbo, J. S., 282
Panchenko, M. P., 443
Pandi, L., 61, 255, 258, 271
Pandya, B. V., 270
Pandya, M. J., 85, 97, 98, 100, 103
Pangilinan, T., 406
Panini, S. R., 177
Panitch, A., 451, 452
Pannekoek, H., 276
Pant, H. C., 175
Papp, A. C., 228
Pappas, G. D., 356
Paramio, J. M., 176
Pardo, J. V., 164, 165, 213
Parise, L. V., 278
Park, D., 275
Park, E. S., 410, 411, 412
Park, J., 151, 152
Park, S. K., 179
Park, T., 172
Parker, T., 449
Parker, T. M., 451

Parker, T. S., 277
Parkes, M., 174
Parkinson, J., 358
Parks, W. C., 440, 442, 443
Park-Snyder, S., 150
Parreira, D. R., 451, 452
Parrott, J. A., 262
Parry, D. A. D., 13, 14, 16, 17, 19, 21, 23, 24, 25, 26, 27, 45, 47, 51, 57, 58, 59, 61, 81, 84, 86, 87, 88, 89, 90, 91, 92, 96, 114, 115, 116, 117, 118, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 131, 132, 133, 134, 135, 149, 150, 153, 157, 159, 169, 170, 181, 256, 307, 346, 350, 353, 356, 360
Parsegian, V. A., 315, 316
Partridge, S. M., 445, 446
Pascoe, L. T., 150
Pascual, J., 206, 207, 211, 218
Pasquali-Ronchetti, I., 416, 447, 448
Passage, E., 440
Passos-Beuno, M. R., 228
Pastan, I., 95
Pasternak, C., 217
Pastore, A., 223, 224
Patchell, V. B., 216
Patel, B., 204
Patel, J. R., 209, 213, 220
Paterlini, M. G., 349
Pattanaik, A., 451
Patterson, C. E., 443
Patterson-Kane, J. C., 360
Patti, J. M., 326
Paulin, D., 166, 174
Pauling, L., 38, 39, 55
Pauptit, R., 225
Pavalko, F. M., 212
Pavanello, R. C., 228
Pawlyk, B., 439, 445
Peachey, A. R., 326, 327, 329
Pearce, M., 213
Pearl, D. K., 228
Pearson, P. L., 228
Pearton, D. J., 169
Pechik, I., 263
Peden, A. A., 178
Pedersen, L. C., 270
Peifer, M., 159
Pekrun, A., 229
Pelin, K., 173
Pelletier, J. N., 58
Pellissier, J. F., 166
Peltonen, L., 427, 442
Peng, Z. Y., 60
Pepe, G., 376
Peppas, N. A., 452
Peranen, J., 357
Pereda, J. M. D., 155, 156
Pereira, L., 406
Perez, P., 176
Pericak-Vance, M. A., 228
Perides, G., 177
Perman, B., 152, 163
Perng, M. D., 174
Perret, S., 317, 319
Perri, T., 451
Perrotta, S., 229
Perry, S. V., 216
Persikov, A. V., 318, 320, 322, 324, 325, 326, 331, 332
Pertz, O., 377
Peter, M. E., 174
Peters, D. J., 162
Peters, H., 406, 411, 413
Peters, J., 45, 53, 54, 55, 60
Peters, R., 410, 411, 412, 415, 418
Petersen, E., 439
Petersen, L. C., 275
Petka, W. A., 99
Petrini, S., 376
Petrof, B. J., 213
Petros, A. M., 225
Petruska, J. A., 346, 347
Petry, F., 330
Peyrol, S., 439, 445
Peyton, D. H., 321, 325
Pfaff, M., 412
Pfander, J. P., 53, 60
Pfuhl, M., 223
Phelps, S. F., 219
Phillips, G. J., 253, 257, 259, 265, 267, 272
Phillips, G. N., 25, 257
Phillips, G. N., Jr., 44, 60, 61
Picaud, S., 213
Picken, L. E. R., 377
Pickeral, O. K., 226
Piedra, J., 160
Pierce, R. A., 414, 415, 442, 443

Pietka, T., 346, 357
Piez, K. A., 303, 314, 346, 350, 356
Pihlajamaa, T., 357
Pihlajaniemi, T., 304
Pillitteri, R. J., 331, 332
Pimenta, A., 179
Pinder, J. C., 221
Pindyck, J., 260
Ping, L. F., 259
Pinto, L., 229
Piontkivska, H., 440, 442
Pique, M. E., 63
Piron, G., 174
Pisa, B. J., 213
Pithawalla, R. B., 114
Pizzo, S. V., 277
Plaas, A. H., 358
Plant, P. W., 260, 261
Plecs, J. J., 55, 87, 89, 96, 97, 100
Plessmann, U., 160
Plow, E. F., 276, 278, 279
Plückthun, A., 58, 59
Podolnikova, N. P., 279
Poirier, G. G., 174
Pokala, N., 214, 215
Pokidysheva, E., 377
Polewski, R., 415, 422
Pollak, M. R., 231, 232
Pollard, T. D., 210
Pollner, R., 377
Pons, F., 213
Ponting, C. P., 165, 210, 227
Poojary, M., 254, 257
Poon, E. P., 165, 166
Pope, B., 214
Popov, V., 98
Porter, G. A., 213
Porter, K. R., 356
Porter, R. M., 174
Portier, M. M., 174
Poshcl, E., 411
Potekhin, S. A., 96, 98
Pothier, B., 230
Potter, J. M., 282
Potter, S., 281
Pottier, S., 221
Pourdeyhimi, B., 451
Powell, B. C., 170, 171
Poy, F., 210, 224, 226
Pozharski, E. V., 218
Pradel, L.-A., 167
Praetzel, S., 170, 171
Prahlad, V., 19, 114, 132, 134, 178, 179, 181
Pramparo, T., 153
Prater, S. M., 152, 153
Pratt, K. P., 257, 259, 280
Prchal, J. T., 229
Prescott, A., 174
Presland, R. B., 168, 169
Presnell, S. R., 70
Prevost, M. C., 174
Price, D., 174
Price, M., 163
Price, T. M., 267
Prins, F., 162
Prior, T. W., 228
Privalov, P. L., 254, 303, 311, 312, 314, 321
Privitera, S., 444
Prockop, D. J., 314, 319, 322, 351, 353, 356, 358
Procyk, R., 269, 272, 273
Prody, C. A., 444
Profumo, A., 269
Proll, V., 159
Propst, F., 156
Pujol, J. P., 388
Pulkkinen, L., 154, 156
Puolakkainen, P., 359
Purkis, P. E., 152, 155, 161
Purslow, P. P., 406, 419, 426
Putnam, E. A., 410, 411, 412, 413
Puustinen, A., 357
Pyke, C., 282
Pytela, R., 154, 156

Q

Qi, Z., 175
Qian, C. P., 416
Qian, R. A., 323
Qian, R. Q., 415, 420
Qiu, D., 179
Qiu, J. J., 230
Quaglino, D., Jr., 447, 448
Que, B. G., 260
Quinlan, R. A., 135, 136, 173, 174

R

Raae, A. J., 219
Raats, C. J., 213
Rabaud, M., 448, 451, 453
Rabindran, S. K., 99
Radmer, R. J., 331
Rafael, J. A., 219, 229
Raghunath, M., 410, 411, 412, 415, 418
Raines, R. T., 302, 315, 321, 322
Rainey, J. K., 355
Raleigh, D. P., 81
Ramachandran, G. N., 18, 302, 303, 307, 311, 312, 314, 319, 321
Ramirez, F., 406, 408, 412, 429, 438, 439
Rampazzo, A., 152
Ramshaw, J. A., 318, 320, 322, 324, 325, 326, 327, 328, 331, 358
Rand, M. D., 269
Randall, J. T., 350
Rando, T. A., 212
Range, A., 171
Rantamaki, T., 427
Rao, H., 412
Rao, M. V., 179
Rao, S., 254, 257
Rapp, G., 366
Rasko, J. E., 453
Rasmussen, M., 306, 323
Raspanti, M., 353, 354, 356, 360
Rau, D. C., 315, 316
Raucher, D., 451
Raurell, I., 160
Ravichandran, K. S., 225
Rawlings, A. V., 169
Ray, P. N., 228
Raymundo, S., 451, 452
Reale, E., 382
Rebecchi, M. J., 224
Récan, D., 228
Rechsteiner, M., 30
Redaelli, A., 360, 366
Redman, C. M., 261
Rees, E., 387
Regan, D. G., 439, 445
Regan, L., 59, 101, 104
Rehemtulla, A., 169, 410, 411, 412
Rehn, M., 304
Reichenbecher, M., 45, 54
Reichenzeller, M., 116, 123
Reid, K. B. M., 304, 306
Reid, R. E., 92, 102
Reimann, J., 156
Reinboth, B., 440
Reinhardt, D. P., 323, 345, 406, 409, 410, 411, 412, 413, 414, 416, 440
Reinholt, F. P., 357, 358, 360, 363
Reipert, S., 155, 156
Reitamo, S., 443
Reitzel, D., 451
Remijn, J. A., 278
Remington, S., 114, 115
Remitz, A., 443
Ren, R., 224, 226
Rendon, A., 213
Renner, C., 377
Rennke, H., 231, 232
Rentschler, S., 224, 226, 227
Repetti, E., 153
Reshetnikova, L., 61
Resing, K. A., 168, 169
Resink, T. J., 448
Resurreccion, E. Z., 174, 176
Revel, J.-P., 265, 350
Reynolds, A. B., 159, 160
Rezniczek, G. A., 155, 156, 163
Ricard-Blum, S., 387
Rice, P., 224
Rich, A., 18, 302, 307, 308, 311
Rich, C. B., 416, 440, 443
Richardson, D., 40, 45, 56, 83, 89
Richardson, J. A., 416, 440
Rickles, F. R., 270
Rief, M., 218
Riemer, D., 114, 134
Ries, A., 327, 379
Rifkin, D. B., 410
Rigby, B. J., 324
Rigler, P., 98
Rijken, D. C., 275, 276
Riley, D. J., 442
Riley, M., 61, 255, 258
Riou, S., 150, 151, 152, 154
Riss, J., 213
Ristori, M. T., 448
Ritty, T. M., 410, 411, 412, 413, 416, 450
Rivero, F., 214, 222
Rixon, M. W., 260
Roark, E. F., 440
Robb, B. W., 412

Robb, S. A., 228, 229
Robert, A., 213
Robert, L., 448
Roberts, E. A., 180
Roberts, H. R., 279
Roberts, R. G., 212, 228, 229
Robertson, S. J., 232
Robinson, M. S., 324
Robinson, P. N., 406, 409, 426, 427
Robson, R. M., 114, 136, 166, 180, 181
Rocco, M., 269
Roche, A., 213
Rock, M. J., 407, 414, 415, 420, 422, 426, 428, 444
Rodger, A., 97, 98
Rodgers, U. R., 448, 452
Rodius, F., 213
Rodriguez-Perez, J. C., 231, 232
Rogers, G. E., 16, 17, 169, 170, 171
Rogers, K. R., 132
Rogus, J. J., 229, 230
Rohlfing, A., 343
Roitbak, T., 162
Rom, E., 387
Romeo, G., 228
Romer, J., 282
Ronchetti, I. P., 448
Rondeau-Mouro, C., 221
Ronziere, M.-C., 344
Rooney, A. P., 443
Rooney, M. M., 278
Roop, D. R., 117, 132
Roper, K., 147
Rosenberg, T., 409
Rosenbloom, J., 319, 412, 415, 416, 440, 441, 442, 443, 444, 445, 448
Rosenheck, S., 152
Roses, A. D., 228
Ross, A., 318
Ross, J., 416
Ross, J., Jr., 440
Rossant, J., 152
Rossi, V., 152
Ross-Murphy, S. B., 272
Roth, R. A., 417, 419
Roth, W., 302
Rothnagel, J. A., 17, 169, 170
Rotter, N., 451
Rousaouen, M., 388
Roustan, C., 216
Rouy, D., 277
Roveri, N., 353
Rovira, A., 451
Ruben, G. C., 377
Rubin, G. M., 226
Rubin, K. M., 284
Rucker, E., 324
Rudall, K. M., 15, 248
Rudee, M. L., 267
Rudnicka, L., 443
Rugg, E. L., 154, 174
Ruggeri, A., 353, 354, 356, 360
Ruggeri, Z. M., 278
Ruggiero, F., 317, 319, 327, 344
Ruhrberg, C., 146, 149, 157, 158
Ruigrok, R. W. H., 60
Ruiz, S., 176
Rujigrok, J. M., 451
Ruoslahti, E., 274
Rutkowski, R., 91
Ruvinov, S. B., 95
Ryadnov, M. G., 98, 100
Ryan, E. A., 268, 273
Rybakova, I. N., 209, 213, 215, 217, 220
Rybarczyk, B. J., 259
Rybicki, A. C., 230
Ryoo, Y., 158

S

Sabatelli, P., 376
Sacca, B., 328, 329
Sackmann, E., 221
Sackmose, S. G., 416
Sado, Y., 379
Sage, E. H., 359
Saharinen, J., 415, 416
Sahlgren, C. M., 175
Sahni, A., 274, 275
Sahni, S. K., 275
Sahr, K., 19
Sahr, K. E., 204
Saidel, L. J., 114
Saido, T. C., 169
Saito, H., 284
Sakai, K., 19, 231
Sakai, L. Y., 406, 408, 409, 410, 411, 412, 413, 415, 416, 417, 428, 429, 439, 440
Sakakibara, S., 319

Sakamoto, H., 412
Sakata, Y., 276
Sakharov, D. V., 275, 276
Salamon, M., 221
Salas, P. J., 175, 177
Salazar, G., 178
Salem, G., 18
Salmikangas, P., 221
Salo, W. L., 114
Salven, P., 204, 209
Salvesen, G. S., 174
Samama, M., 267
Samatey, F. A., 63
Sample, J., 181
Sanada, K., 179
San Antonio, J. D., 306, 326, 327
Sanchez, M., 169
Sandberg, L. B., 438, 445, 449
Sandell, L. J., 307
Sanders, J. E., 360
Sanes, J. R., 166
Sanger, J. M., 164, 213
Sanger, J. W., 163, 164, 213
Sanguineti, C., 406, 408, 439
Santoro, S. A., 274
Santos, G. F., 213
Santos, M., 176
Saraste, J. T., 214, 221
Saraste, M., 204, 206, 207, 209, 211, 214, 218, 219, 220, 221, 223, 224, 225
Sargent, D., 56, 88
Sarkar, S. K., 307, 316, 318
Sarria, A. J., 177
Sasaki, M., 13, 17, 159
Sasaki, N., 365, 366
Sasaki, T., 269, 416, 440
Sasaoka, T., 166
Satsangi, J., 174
Sattler, M., 225
Sauer, R. T., 54, 88
Saund, A. K., 92, 102
Saurat, J. H., 150, 151, 152, 154
Sauter, H., 151
Savidge, G. F., 262
Savontaus, M.-L., 228
Savvides, P., 229
Sawada, N., 222
Sawaguchi, S., 353
Sawamura, D., 149, 153
Saxe, D., 260
Scacheri, P. C., 376
Scanu, A. M., 277
Scarlata, S., 224
Scarpa, A. L., 204
Scarpa, E., 19
Schaapveld, R. Q., 149, 152, 155
Schachar, N. S., 361
Schachat, F., 219
Schaffer, F., 229, 230
Scharrer, I., 262, 281
Schaub, T., 412
Scheffers, M. S., 162
Scheffzek, K., 209, 218, 219, 220, 221
Scheiner, T., 282
Scheraga, H. A., 264, 315, 348, 349
Scherbarth, A., 177
Scherer, S. E., 427
Scherpf, S., 228
Schickling, O., 174
Schittny, J. C., 380
Schlaepfer, D. D., 222
Schleicher, M., 222
Schlessinger, J., 224, 225
Schlichting, I., 19, 20, 21, 22, 124
Schlosser, A., 416
Schlunegger, M. P., 215
Schmelz, M., 149
Schmitt, D., 169
Schmitt, F. O., 355
Schmittner, M., 166
Schnarr, N. A., 89
Schneider, J. P., 82, 93
Schneider, V., 228
Schnittler, H., 232
Schnolzer, N., 160
Schreiber, B. M., 444
Schröder, R., 156
Schroter, W., 229, 230
Schubert, M., 416
Schulthess, T., 21, 54, 59, 84, 124, 309
Schulze, K., 151, 160
Schwabe, J. W. R., 29
Schwaeble, W., 416
Schwartz, R. S., 230
Schwarz, M. A., 149
Schwarzbauer, J. E., 412
Schwarze, U., 331, 332
Schweizer, J., 170, 171
Scopa, A., 450
Scott, C., 168

Scott, G. A., 161
Scott, I. R., 169
Scott, J. E., 357, 363
Scott, P. G., 347, 349
Scrutton, M. C., 272
Sebbag, M., 169
Sechler, J. L., 412
Sedra, M. S., 228
Segrade, F., 414, 415
Segrelles, C., 176
Seifert, G. J., 156
Seitz, M., 449
Seki, T., 379
Sellin, L. K., 161, 162
Senga, K., 363
Senior, R. M., 448
Seo, J., 44, 45, 48, 51, 53
Seo, S., 163
Serafini-Fracassini, A., 447
Sernett, S. W., 114, 166, 180
Serre, G., 169
Sessions, R., 44, 100
Sethson, I., 223
Settler, A., 151, 160
Setton, L. A., 451
Setzer, S. V., 151
Sevcik, J., 155
Sevilla, L. M., 158
Sewry, C., 173, 228
Seyama, Y., 450
Shackleton, D. R., 358
Shafer, J. A., 250, 264
Shah, G. A., 273
Shah, J. V., 179
Shah, N. K., 328
Shainoff, J. R., 264, 265, 267, 269, 280
Shalaby, S. W., 451, 452
Shalev, O., 229, 230
Shanmugasundararaj, S., 379
Shapiro, J. R., 343, 346
Shapiro, S. D., 426, 428
Sharma, M., 328
Sharma, P., 175
Sharma, V. A., 96
Shaw, G., 225
Shaw, J. M., 327, 328, 330
Shaw, R., 225
Shea, T. B., 179
Sheiba, L., 449
Shekhel, J. J., 61
Shen, W., 172, 214, 215
Shenoy, N., 331
Sheppard, P., 442
Sher, B., 261
Sheren, S. B., 350
Sherratt, M. J., 381, 382, 387, 406, 411, 414, 417, 418, 419, 420, 422, 423, 425, 426, 428, 439, 440
Shete, S. S., 427
Shevach, E. M., 152
Shi, J., 161
Shibata, H., 221
Shields, P. P., 264
Shigeno, M., 353
Shigesada, K., 172
Shikata, K., 304
Shimakawa, M., 221
Shimizu, H., 147, 154, 156
Shimizu, N., 19
Shimizu, T., 213
Shimomura, I., 304
Shin, C. K., 256, 259
Shin, D. S., 63
Shinkai, H., 159
Shipway, J. M., 61
Shirai, H., 306
Shirakawa, S., 280
Shirato, I., 231
Shirinsky, V. P., 448
Shishibori, T., 416
Sholders, A. J., 56, 60, 92, 105
Shorny, S., 156
Shrager, J. B., 213
Shtutman, M., 159
Shu, T., 179
Shudo, K., 319
Shue, C., 226
Shukunami, N., 365
Shulman, N. R., 275
Shumaker, D. K., 144
Shuman, M. A., 270
Shum-Tim, D., 451
Shuttleworth, A., 382, 386, 388, 406, 411, 414, 415, 416, 417, 418, 419, 420, 422, 426
Shuttleworth, C. A., 381, 382, 386, 387, 388, 406, 407, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 422, 423, 425, 426, 428, 439, 440, 444
Sia, S. K., 95

Sibley, R. K., 174
Siebenlist, K. R., 250, 264, 267, 268, 270
Siebold, B., 323, 379, 380
Sigler, P. B., 224, 225
Sigrist, C. J., 13
Sikaneta, T., 162
Sikorski, A. F., 221
Siliciano, J. D., 164, 213
Sillivan, K. A., 439, 445
Sillman, A. L., 204, 207, 210
Silveira, A., 272
Silver, F. H., 349
Simionati, B., 152
Simon, D. I., 281
Simon, M., 149, 157, 169
Simon, M. I., 260
Simon, S. R., 273
Simon-Lukasik, K. V., 318
Simpson-Haidaris, P. J., 250, 252, 255, 259, 261
Singer, J., 387
Singh, M., 47, 58, 63
Singh, N., 379
Sinha, S., 415, 416
Siracusa, L. D., 357, 358, 418
Sittinger, M., 451
Sixma, J. J., 278
Sjostrom, M., 448
Sjuve, R., 166
Skaer, R. J., 377
Skalli, O., 154
Skehel, J. J., 99
Skerrow, C. J., 149
Skinner, M. K., 262
Skupski, M. P., 226
Slavin, R. E., 355
Slayter, H. S., 253, 382
Sloan, C., 445, 446
Small, J. V., 155, 165, 166, 171, 172
Small, K., 281
Smallridge, R. S., 408
Smejkal, G. B., 264
Smillie, L. B., 90
Smirnov, V. N., 448
Smith, B. O., 222
Smith, E. A., 151, 160
Smith, F. J. D., 154, 156
Smith, F. J. H., 156
Smith, J. J., 80
Smith, J. W., 350
Smith, K. D., 359
Smith, M. S., 412
Smith, T. A., 117, 120, 121, 122, 123, 124, 125, 126, 127, 128, 131, 132, 307
Smola, H., 163
Smoyer, W. E., 231
Snowhill, P. B., 349
Snyder, P. J., 228
Soetikno, R. M., 174, 175
Sokol, S. Y., 161
Solomon, E., 440
Soloviev, D. A., 279
Solovjov, D. A., 279
Sombret, B., 447
Sommer, P., 439, 445
Soncini, M., 360, 366
Sonenshein, G. E., 443
Song, M. J., 363
Sonnenberg, A., 147, 149, 150, 151, 152, 153, 154, 155, 215, 217
Sonnhammer, E. L., 226
Sonnichsen, F. D., 90, 103
Sorensen, L. K., 438, 448
Soria, C., 267, 275
Soria, J., 267, 275, 280
Soskel, N. T., 438
Soslowsky, L. J., 358
Southan, C., 283
Spann, T. P., 144
Sparrow, L. G., 21, 23, 114
Spazierer, D., 156, 159, 163
Speed, T., 47
Speer, A., 227, 228
Speicher, D. W., 19, 204, 206, 218
Spencer, J. A., 439, 445
Spina, M., 447
Spitzer, R. H., 114
Spitzer, S., 262, 281
Spitznagel, L., 357
Spolski, R. J., 260
Spooner, G. M., 97, 98
Sporn, L. A., 262, 274, 278
Spottke, B., 343
Spraggon, G., 255, 257, 258, 259, 260, 269, 270, 276, 283
Spruit, L., 162
Spurr, N., 19
Spurr, N. K., 213
Squire, J. M., 18, 25, 27, 303, 377, 382, 383, 385, 388, 389, 390, 391, 392, 393

Stafford, W. d., 91
Standeven, K. F., 282
Stanley, J. R., 149, 152, 158
Stappenbeck, T. S., 150, 151
Starcher, B. C., 416, 439, 440, 442, 445, 453
Stauffacher, C. V., 253, 254, 257, 267
Stedman, H. H., 213
Stedronsky, E. R., 451
Steel, E., 59
Stegh, A. H., 174
Stehle, T., 162
Steigemann, W., 323
Steinberg, L. S., 228
Steinert, P. M., 17, 19, 23, 59, 114, 115, 116, 117, 118, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 131, 132, 133, 134, 135, 149, 150, 151, 157, 158, 168, 169, 170, 181
Steinmann, B., 427
Steinmetz, M. O., 59, 84, 100, 126
Steitz, T. A., 61
Stenkamp, R. E., 257, 259, 270
Stepanova, O. V., 448
Stephan, S., 387
Steplewski, A., 324
Stetefeld, J., 21, 54, 82, 86, 124, 309
Stetter, K.-O., 53, 60
Stetzkowski-Marden, F., 167
Steven, A. C., 30, 114, 117, 121, 127, 132, 157, 158, 169
Stevens, H. P., 152, 161
Stevens, M. M., 101
Stewart, F., 60
Stewart, M., 14, 24, 25, 26, 45, 57, 60, 90, 126, 149, 154, 388, 389
Stewart, R. J., 99
Stinchombe, T. E., 279
Stirling, Y., 272
Stöcher, M., 155
Stock, A., 59, 84
Stock, J., 47, 49, 86, 90, 91
Stock, U. A., 451
Stockinger, A., 154
Stoddard, B. L., 81
Stolle, C. A., 442
Stone, P. J., 440, 444
Stoneking, M., 440
Stoneman, R. G., 85, 97, 98, 100, 103
Stpierre, S. A., 92, 102
Stradal, T., 214
Strand, J., 61
Strasser, G., 153
Stratowa, C., 149, 154
Straub, V., 212, 213
Streeten, B. W., 416, 419, 423
Strelkov, S. V., 19, 20, 21, 22, 44, 45, 48, 53, 82, 86, 121, 123, 124, 125, 126, 127, 128, 129, 131, 132
Stringer, H. A., 276
Stromer, M. H., 114
Strong, D. D., 267, 279
Strong, L. E., 250
Struminger, C., 440
Stryer, L., 265
Stuart, D. I., 408, 409, 412, 418, 420, 423, 425
Stubbs, M. T., 263
Stull, J. T., 166
Stultz, C. M., 318
Stumpp, M. T., 59
Stumptner, C., 174
Styers, M. L., 178
Su, J., 19
Su, J. Y., 84
Subramony, S. H., 376
Sudol, M., 22, 210, 224, 226, 227
Suehiro, K., 278
Suenson, E., 275
Suga, Y., 169
Sugimoto, M., 379
Sugita, Y., 153
Sugo, T., 279
Sugrue, S. P., 161
Suh, T. T., 281
Sui, S. F., 221
Sukhatme, V. P., 162
Sullivan, B. M., 162
Sullivan, C. E., 307, 316, 318
Sullivan, N., 166
Sum, H., 221
Sun, D., 153, 155, 179
Sun, T. T., 169, 170
Sun, X., 304
Sunada, Y., 219
Sundaramoorthy, M., 330, 379
SundarRaj, N., 344
Sunde, M., 97, 98
Sutherland, G. R., 416
Sutherland-Smith, A. J., 214, 217
Suthers, G. K., 213
Sutmuller, M., 386, 387

Suyama, K., 445
Suzigan, S., 451, 452
Suzuki, E., 18, 126, 127, 307, 308, 314, 316, 321, 322, 350, 353
Suzuki, K., 169
Suzuki, N., 103
Suzuki, R., 363
Suzuki, T., 163
Svensson, L., 358, 360
Svensson, O., 357, 363
Svitkina, T. M., 154, 155, 163
Swartz, K. R., 451, 452
Swedo, J. L., 355
Swee, M. H., 442, 443
Sweeney, H. L., 58, 213
Sweeney, S. M., 326, 327
Sykes, B., 406
Sylvestre, P., 306
Szabo, Z., 440
Szilak, L., 89
Sztul, E., 177

T

Tachibana, K., 163
Tainer, J. A., 63
Taipale, J., 415, 416
Tajima, S., 416
Takada, Y., 416, 440
Takahara, Y., 327
Takahashi, A., 174
Takahashi, K., 171
Takahashi, S., 363
Takahashi, T., 175
Takakuwa, H., 181
Takanaga, H., 221
Takata, R. I., 228
Takayanagi, M., 303, 304, 329
Takeda, Y., 262
Takenawa, T., 222
Takeo, N., 13, 17, 159
Takeuchi, M., 445
Talbot, J. A., 84, 98
Talluri, S., 224
Talmage, K., 282
Tamai, K., 443
Tamburro, A. M., 438, 442, 447, 450, 451
Tanaka, T., 223
Tang, D., 175
Tang, J., 215, 216, 222
Tanji, K., 213
Tao, T., 172
Tarbouriech, N., 60
Tarcsa, E., 135, 169, 170
Tashiro, K., 416, 440
Taylor, D. G., 448
Taylor, D. W., 215, 216, 219, 222
Taylor, H. L., 250
Taylor, J. M., 67
Taylor, K. A., 215, 216, 219, 222
Taylor, K. M., 322
Taylor, R. K., 63
Teach, J., 451, 452
Teisner, B., 416
Teller, D. C., 168, 169, 257, 259, 270
Terpstra, A. J., 443
Textor, G., 451
Thakore, N., 376
Than, M. E., 379
The, C., 304, 306
Theisen, M., 317, 319
Therkildsen, L., 267
Thiagarajan, P., 277
Thiel, G., 226
Thiene, G., 152
Thiyagarajan, P., 99
Thomas, G. H., 206, 208, 410, 412, 413
Thomas, R. M., 56, 88
Thomlinson, A. M., 357, 363
Thompson, J., 224
Thompson, J. B., 356
Thompson, T. G., 165
Thornell, L.-E., 164, 165
Thorpe, J. R., 97, 98
Thorsen, S., 275
Thulin, C., 169
Tidor, B., 55, 87, 89, 96, 97
Tiecke, F., 409
Tiedemann, K., 323, 345, 406, 411, 412, 413, 416
Tiller, G. E., 442
Timmins, P. A., 348
Timpl, R., 377, 379, 380, 382, 386, 412, 416, 440
Tinsley, J. M., 210, 213, 214, 222, 224, 226, 227
Tint, I. S., 163
Tirrell, D. A., 99, 101, 448, 451, 452
Tisdale, C., 410, 412, 413, 445, 446

Titeux, M., 166
Tixier, J. M., 448
Tkachuk, V. A., 448
Tobacman, L. S., 61
Tochetti, D., 350
Todd, P., 379
Toilola, D. M., 177
Tome, F., 174
Tomino, Y., 231
Tommasi di Vagnano, A., 226
Tonchev, T., 47
Tonegawa, S., 284
Tong, H. Q., 231, 232
Tooney, N. M., 250, 257
Toonkool, P., 439, 445
Torbet, J., 265, 269, 273
Torchia, D. A., 117, 307, 316, 318
Torgersen, D., 327, 328, 330
Tornoe, I., 416
Tornvall, P., 272
Toselli, P., 444
Toshimori, M., 221
Towbin, J. A., 152
Townsend, R. R., 256
Toyoshima, T., 416
Tran, H., 274
Tran, V. T., 427
Trask, B. C., 414, 415, 416
Trask, T. M., 412, 413, 414, 416
Traub, P., 155, 177
Traub, U., 177
Traub, W., 18, 303, 307, 314
Trave, G., 223
Trelastd, R. L., 360
Trevisan, S., 221
Trinnaman, B., 151, 161
Tripet, B., 59, 89
Trivedi, M., 174
Trong, I. L., 257, 259
Troyanovsky, R. B., 151
Troyanovsky, S. M., 151
Trus, B. L., 127, 350
Trybus, K. M., 58
Tryggvason, K., 330
Tsai, L. H., 179
Tschodrich-Rotter, M., 410, 411, 412
Tse, W. T., 230, 231
Tseng, H. C., 179
Tsiokas, L., 162
Tsipouras, P., 416
Tsukita, S., 19, 161
Tsurupa, G., 274, 276, 277
Tsuruta, D., 147, 156, 163
Tu, Y. H., 172
Tuckwell, D. S., 326, 327, 329
Tufty, R. M., 222
Tukalo, M. A., 85
Turci, M., 269
Turner, G., 219
Turner, M. W., 330
Tuszynski, G. P., 274
Tzivion, G., 177

U

Uchiyama, S., 322
Ueda, J., 382, 386
Ueda, M., 363
Ueki, Y., 379
Uematsu, J., 154
Uesugi, S., 306, 327
Ugarova, T. P., 254, 265, 279
Uhrin, D., 222
Uitdehaag, J., 327, 328, 330
Uitto, J., 149, 153, 154, 156, 157, 158, 440, 441, 442, 443
Umeda, H., 445
Underwood, R. A., 169
Unsold, C., 415
Uragami, T., 452
Urban, Z., 450
Urbanikova, L., 155
Urbanke, C., 343
Urry, D. W., 444, 445, 447, 449, 450, 451, 452
Urtizberea, J. A., 173
Ushiki, T., 353

V

Vacanti, J. P., 451
Vainchenker, W., 262
Vainzof, M., 228
Vaittinen, S., 175
Valdre, U., 447, 448
Vale, R. D., 179
Valle, G., 221
Vallee, R. B., 178

Valmu, L., 357
Vanacore, R. M., 379
van den, I. P., 174
van den Born, J., 213
van der Bent, P., 162
Vanderhyden, B. C., 232
van der Rest, M., 387
van der Ven, P. F., 156
Vandyke, M., 86, 90, 91
Van Dyke, S., 47, 49
van Heel, D., 174
van Hemel, B. M., 278
van Hest, J. C. M., 448
VanLoock, M. S., 215, 217
van Ommen, G. J. B., 228
van Paassen, H. M. B., 228
van Roy, F., 151, 160
van't Veer, C., 269
van Wilpe, S., 155
van Zonneveld, A. J., 276
Varani, G., 81
Vardaxis, N. J., 451
Varga, Z., 448
Varret, M., 429
Vasiliev, V., 98
Vasioukhin, V., 152, 161
Vasseur, C., 229
Vassiliadis, J. N., 229, 230, 231
Vaughan, L., 346
Vazina, A. A., 98
Vedanarayanan, V., 376
Veerapandian, L., 61, 255, 258, 259, 260, 269
Veerman, H., 276
Vega-Colón, I., 412
Veillard, A., 388
Veis, A., 347, 348
Veklich, Y. I., 254, 267
Vena, P., 360, 366
Venturoni, M., 356
Venugopal, M. G., 308, 309, 312, 314, 315, 316, 318
Verdetti, J., 448
Verdini, A. S., 98
Verkhovsky, A. B., 155
Verna, A., 448, 451
Vesely, I., 450
Vesentini, S., 360, 362, 366
Vicart, P., 174
Vicente, V., 278
Viel, A., 206, 208, 225
Vilaire, G., 277, 278
Villafranca, J. J., 95
Villard, V., 98
Vincent, M., 221
Vincent, P. A., 147, 149
Vincenzi, D., 416
Vindigni, A., 263
Vinson, C., 57, 58, 89, 95
Vinzens, U., 151, 160
Virata, M. L. A., 149, 150, 151
Visse, R., 319
Vitagliano, L., 309, 314, 315, 316, 319, 322, 348
Vlcek, S., 144, 145
Voit, T., 213, 227, 228
Volderviszt, F., 63
Volkmann, N., 25, 61, 215, 255, 258, 270
Vollbrandt, T., 406, 411, 413
Vollrath, F., 351
Volpin, D., 440, 447, 448
Von Der Mark, H., 382
von Hippel, P. F., 303, 314
von Hippel, P. H., 312
Vosshall, L. B., 226
Vrhovski, B., 439, 440, 444, 445, 447, 449, 453
Vrielink, A., 177

W

Wachi, H., 412, 443, 450
Wada, H., 280
Wada, Y., 306, 327
Waddell, M. J., 322
Wade, W. S., 225
Wadia, Y., 451, 452
Wagberg, F., 439
Wagner, G., 224
Wagner, R. M., 149, 150
Wagschal, K., 59, 89
Waite, E., 85, 97, 98, 100, 103
Wakita, M., 365
Walke, S., 214, 215, 216
Walker, J. B., 276
Wall, J., 253, 267, 268
Wallace, R. N., 419, 423
Wallach, J., 448
Wallgren-Pettersson, C., 173

Wallis, D. D., 410, 412, 413
Wallis, O., 451
Wallis, R., 327, 328, 330
Walrafen, G. E., 315, 316
Walsh, F. S., 213, 219
Walsh, K. A., 169
Walsh, M. P., 214
Walshaw, J., 48, 61, 84, 86, 87, 88, 89, 90, 96, 97, 101, 102, 105
Walter, U., 19, 20, 21, 22, 124
Walton, K. W., 355
Walz, D. A., 274
Walz, G., 161, 162
Wandinger-Ness, A., 162
Wang, C., 99
Wang, C. C., 30, 101
Wang, C. L., 172
Wang, D. S., 225
Wang, H., 127, 132
Wang, J. H., 175
Wang, K., 172
Wang, L., 179
Wang, N., 163
Wang, P., 171, 172
Wang, Q., 330
Wang, W., 80, 276
Wang, X., 59, 311
Wangh, L. J., 260
Wanner, G., 377
Ward, C., 162
Ward, D. C., 230
Ware, C. F., 174
Ware, R. E., 231
Warren, S. G., 257, 270
Warren, S. L., 209, 220, 231
Warren, V., 216
Wasenius, V. M., 204, 209
Wasiak, S., 153
Wasmuth, J. J., 416
Watanabe, M., 363
Watanabe, N., 304
Watkins, P. C., 230
Watkins, S. C., 165, 166
Watson, R. E., 426
Watt, F. M., 146, 149, 157, 158
Wattanasirichaigoon, D., 173
Waugh, R. E., 218
Way, M., 214, 215, 216
Weaver, A., 19
Weaver, A. C., 152
Weaver, D. C., 167
Weber, K., 114, 117, 134, 160, 166, 168, 172, 180
Wedig, T., 123, 125, 132
Weeds, A. G., 214
Wei, S. M., 447
Weinfield, M., 277
Weinhold, F., 322
Weinreb, C. K., 100
Weir, A., 212, 227
Weis, M. A., 346, 357
Weis, W. I., 150
Weisel, J. W., 250, 253, 254, 256, 257, 259, 260, 265, 267, 268, 270, 271, 272, 273, 275, 276, 277, 278, 280, 282
Weis-Fogh, T., 449
Weiss, A. S., 57, 407, 414, 415, 422, 439, 440, 441, 444, 445, 447, 448, 449, 452, 453
Weissman, N., 445
Weitzer, G., 149, 154
Weksler, B., 156, 163
Wells, D. J., 219
Wells, K. E., 219
Welsh, M. J., 231
Wen, C. K., 356
Wendel, D. P., 448
Wendt, T., 219
Wenglen, C., 382, 386
Werkmeister, J. A., 327, 358
Werneck, C. C., 416
Werner, J. M., 408, 423
Wess, L., 350, 351, 352
Wess, T. J., 348, 349, 350, 351, 352, 353, 354, 356, 406, 411, 414, 417, 419, 420, 426
Wessel, G. M., 210
Wessel, H., 344
Westgate, G. E., 152
Whaley, P. D., 262
Wheeler, G. N., 155
Whitaker, J. A., 104
Whitby, F. G., 30, 60, 61
White, J., 358
White, R., 96
White, S. W., 348
Whiteman, P., 408, 409
Whiting, K., 169
Whitledge, J. R., 451
Whittaker, S. P., 382, 416, 417, 426
Whittock, N., 152, 161
Wiberg, C., 382, 386

Wiche, G., 149, 150, 154, 155, 156, 157, 159, 163, 167
Wick, G., 382
Wiedemann, H., 380, 382, 386
Wiederschain, D., 451
Wierenga, R., 225
Wieschaus, E., 159
Wieslander, J., 379
Wight, T. N., 359
Wilbourn, B., 262
Wiley, D. C., 61, 99
Wilkens, S. J., 322
Williams, D. S., 180
Williams, M. W., 213
Williams, R. C., 253
Williamson, M. P., 226
Williamson, R., 228
Willing, M. C., 427
Willins, M. J., 356
Willis, A. C., 416
Wilmanns, M., 224, 225
Wilson, A. M., 360
Wilson, D. B., 47
Wilson, I. A., 61
Wilson, K. L., 144, 145
Wilson, N. J., 174
Wilson, P. D., 162
Wilson, R., 412, 413
Wilson, S. E., 448, 451
Wilton, S. D., 16, 17, 173
Wiman, B., 275
Winder, S. J., 22, 28, 207, 208, 210, 212, 213, 214, 215, 216, 219, 220, 222, 224, 226, 227
Winkelmann, D. A., 360
Winkelmann, J., 19, 204, 209, 211, 213
Winkler, F. K., 100
Winlove, C. P., 221
Winnard, A. V., 228
Winograd, E., 206, 210
Winokur, S. T., 221
Winter, H., 170, 171
Winter, S. S., 231
Winterhalter, K. H., 346
Wirtz, D., 99, 168, 180
Wisniewski, J., 99
Witke, W., 222
Witt, C. C., 173
Witte, D. P., 279
Wittinghofer, A., 19, 20, 21, 22, 124
Wolf, E., 19, 20, 21, 22, 47, 124
Wolf, J. L., 230
Wolf, Y. I., 80
Wolfe, B. L., 443
Wolfe, L. C., 231
Wolfenstein, T. C., 260
Wollman, R., 153
Wollmann, R., 152, 163
Wong, P. C., 114, 128, 144, 173, 174, 177
Wong, S., 217
Woodhead, J. L., 280
Woodhead, J. W., 276
Woodhead-Galloway, J., 346, 351, 356
Woodhouse, K. A., 439, 447, 451, 453
Woodruff, S., 360
Woods, C. M., 230
Woods, H. L., 104
Woods, T. C., 451
Woolfson, D. N., 47, 48, 55, 56, 61, 81, 84, 85, 86, 87, 88, 89, 90, 91, 92, 95, 96, 97, 98, 99, 100, 101, 102, 103, 105
Woolley, D. E., 426
Wortman, J. R., 226
Wrenn, D. S., 448
Wright, D. M., 419
Wright, G. M., 453
Wright, J., 172
Wright, T. L., 174
Wrogemann, K., 228
Wu, C., 99
Wu, G. J., 174
Wu, J. J., 346, 357, 382
Wu, K. C., 19, 59, 114, 122, 124, 128, 129, 131, 132, 134, 181
Wu, M., 443
Wu, W. J., 440

X

Xia, C. H., 180
Xia, H., 221
Xiao, K., 147, 149
Xie, H., 451, 452
Xie, R. L., 100
Xie, X., 60
Xiong, Y., 59
Xu, G. M., 162
Xu, J., 449, 451, 452
Xu, W., 261

Xu, X., 283
Xu, Y., 306, 311
Xu, Z., 114
Xue, Z. G., 166

Y

Yabe, J. T., 179
Yadav, S. P., 279
Yaen, C. H., 279
Yakovlev, S., 255, 258, 276, 279
Yakubenko, V. P., 279
Yamada, H., 213
Yamada, S., 168, 180
Yamada, Y., 306
Yamakawa, H., 153
Yamamoto, M., 63
Yamamoto, S., 353
Yamamoto, T., 363, 365
Yamaoka, T., 451, 452
Yamashita, K., 416
Yamazaki, M., 169
Yamazumi, K., 259, 280
Yanagida, Y., 452
Yanagisawa, H., 416, 439, 440, 445
Yanagisawa, M., 416, 440
Yandell, M. D., 226
Yang, B., 227
Yang, H.-Y., 154
Yang, J.-M., 59, 124, 131
Yang, W.-H., 315, 316
Yang, Y., 144, 152, 153
Yang, Z., 61, 179, 255, 258, 271
Yao, J., 232
Yao, Y., 30
Yaoita, H., 156
Yaremchuk, A. D., 85
Yau, S. C., 229
Yawata, Y., 230
Yayon, A., 386, 387
Yazaki, K., 98
Yazaki, Y., 306, 327
Yeager, M., 63
Yeates, T. O., 99
Yee, R. W., 312
Yee, V. C., 257, 259, 270, 279
Yeh, H., 440, 441, 442
Yeung, T., 59
Yi, S. J., 230, 231
Yin, M., 152
Ylanne, J., 209, 217, 218, 219, 220, 221
Ylostalo, J., 357
Yoder, N. C., 89
Yonath, A., 307
Yonekura, K., 63
Yong, S. L., 442
Yoon, H. S., 225
Yoon, K. H., 180
Yoon, K. J., 156, 163
Yoon, M., 178, 179, 180
Yorifuji, H., 156
Yoshida, M., 227
Yoshida, N., 280, 319
Yoshida, T., 98
Yoshikawa, T., 181
Yoshioka, H., 13, 17, 159
Yoshiura, K., 429
Young, B. B., 344
Young, P., 209, 217, 218, 219, 220, 221
Young, P. E., 307, 318
Young, R. D., 343
Yu, Q.-C., 152, 153, 163
Yu, S. Y., 261, 284
Yuan, A., 179
Yuan, J., 136
Yuan, X., 409, 410
Yue, B. Y., 382, 386
Yurchenco, P. D., 209, 220, 377, 379, 380
Yuspa, S. H., 152

Z

Zagari, A., 309, 314, 315, 316, 319, 322, 348
Zamir, E., 163
Zarrinpar, A., 226
Zastrow, M. S., 144, 145
Zatloukal, K., 174
Zatz, M., 228
Zaucke, F., 357
Zhang, H., 406, 408, 439
Zhang, J., 226
Zhang, J. Z., 261
Zhang, L., 45, 279
Zhang, M., 223, 442, 443
Zhang, P., 224

Zhang, Q., 162
Zhang, R., 210, 224, 226
Zhang, R. Z., 376
Zhang, T., 19, 20, 43, 44, 56, 60, 87, 89, 90, 91, 93, 96
Zhang, X., 452
Zhang, Y., 440, 442
Zhao, Q., 226
Zhao, Y., 439, 445
Zhen, Y. Y., 161
Zheng, M., 152, 153
Zheng, X. H., 226
Zhou, J., 379
Zhou, M., 99
Zhou, M. M., 225
Zhou, N. E., 58, 102, 103
Zhou, P. H., 344
Zhou, X., 451
Zhu, D., 330
Zhu, E., 153
Zhu, X., 175
Zhurinsky, J., 159
Ziemiecki, A., 171
Ziese, U., 351, 354, 417, 418, 419, 420, 422, 423
Zimbello, R., 152
Zimbelmann, R., 123, 132
Zimenkov, Y., 99
Zimmerman, L. B., 128
Zimmermann, D. R., 416
Zinkham, W. H., 218
Zizak, I., 366
Zörer, M., 156
Zu, Y., 172
Zubrzycka-Gaarn, E. E., 219, 228
Zuerbig, S., 213
Zuffardi, O., 153
Zuo, J., 439, 445
Zurdo, J., 100
Zychband, E. I., 360

SUBJECT INDEX

A

α-helix
  amphipathic coiled coil design and, 82–83
  coiled coil structures and, 38–39
ABC design
  coiled coil design and, 95
ABD. *See* Actin-binding proteins
ABS. *See* Actin-binding sites
Acidic glycoprotein (AGP)
  collagen fibrils and, 359
Acidic residues
  IF chains and, 117, 121, 122, 125, 128, 132
  sequence repeats and, 12
Actin rich cortex
  IFAP structure/function and, 162
Actin-binding (ABD) proteins
  IFAP structure/function and, 146, 155, 156, 158, 167, 176–177
  spectrin superfamily and, 203, 208–210, 213–217, 219, 220, 230
Actin-binding sites (ABS)
  spectrin superfamily and, 215–217, 222
Actinins
  fibrous proteins and, 1, 5, 10
  spectrin superfamily and, 203, 211, 215, 216, 217, 219, 220–221, 222, 232
Actins
  IFAP structure/function and, 146, 156, 171–173
  sequence repeats and, 14
Actin-stress fibers
  spectrin superfamily and, 204, 211
Afibrinogenemias
  fibrinogen/fibrin and, 281
AFM. *See* Atomic force microscopy
Aggregation
  hexagonal aggregation and, 376, 389, 390, 391
  IF chains and, 122
  linear/lateral aggregation and, 376, 380, 388, 396, 399
Aging
  elastin and, 451
  fibrillin microfibrils and, 426
  network-forming collagens and, 376
AGP. *See* Acidic glycoprotein
Albumin
  fibrinogen/fibrin and, 273–274
Allysine
  elastin and, 445–446
Alport syndrome
  network-forming collagens and, 379
Amyloids
  fibrous proteins and, 2
Ankyrin
  IFAP structure/function and, 166, 167–168
  spectrin superfamily and, 211, 220, 225, 230, 231
Antiparallel structures
  coiled coil design and, 101–105
APC tumor suppressor protein probe
  coiled coil design and, 96
Apolar residues
  IF chains and, 117, 118, 121, 124, 126, 129
  sequence repeats and, 12, 15, 19–21
Apolipoproteins
  coiled coil structures and, 67, 69, 70
Apoptosis
  IFAP structure/function and, 174–175
Armadillo proteins
  IFAP structure/function and, 159–161
Assembly. *See also* Coassembly
  of elastin, 444–445
  of fibrillin microfibrils, 412–417
  fibrinogen/fibrin and, 261–262
  of IF chains, 131–132, 136

Assembly. *See also* Coassembly (*cont.*)
of network-forming collagens, 379–380, 384, 385, 389
Atomic force microscopy (AFM)
fibrillin microfibrils and, 411, 417, 420
Axial packing
in collagen fibrils, 345–346, 347

B

β-sheets
coiled coil design and, 83
IF chains and, 120, 133
Basement membranes
network-forming collagens and, 376, 379, 386
Basic residues
IF chains and, 117, 121, 122, 125, 126, 132
sequence repeats and, 12
Beads-on-a-string
fibrillin microfibrils and, 405, 406, 407
Becker muscular dystrophy (BMD)
spectrin superfamily and, 212–213, 219, 227–229
Belt-and-braces designs
coiled coil design and, 99–100
Bethlem myopathy
network-forming collagens and, 376
Binding pockets
fibrinogen/fibrin and, 269
Biological properties
of elastin, 448
Biomaterials
elastin and, 448, 450–453
Biosynthesis
of fibrinogen/fibrin, 260–262
BMD. *See* Becker muscular dystrophy
BP. *See* Integrin binding protein
BP230 isoforms
IFAP structure/function and, 152–154
Buried polar groups
coiled coil design and, 91

C

Calcium binding/EF hands
fibrinogen/fibrin and, 247, 257, 259–260, 269, 270, 277, 284
fibrous proteins and, 5, 27–28
IFAP structure/function and, 168, 169, 170
sequence repeats and, 14–15, 27–28
spectrin superfamily and, 203, 208, 209–210, 222–224, 226
Calcium chelation
in fibrillin microfibrils, 419
Calcium-binding epidermal growth factor (cbEGF)
elastin and, 439
in fibrillin microfibrils, 407, 408–409, 410, 411, 415, 420, 427, 428
Calponin/CH domains
IFAP structure/function and, 153, 155, 156, 171–172
Carcinomas
fibrinogen/fibrin and, 261, 282, 283
cbEGF. *See* Calcium-binding epidermal growth factor
CD. *See* Circular dichroism studies
Cell adhesion
fibrinogen/fibrin and, 274, 278, 285
spectrin superfamily proteins and, 211
Cell cycles/kinases
IFAP structure/function and, 175–176
Cell homeostasis
IFAP structure/function and, 144, 181
Cell process control
IFAP structure/function and, 181–182
Cell surface interactions
IFAP structure/function and, 159–168
CH domains
IFAP structure/function and, 153, 155, 156, 171–172
spectrin superfamily and, 204, 208, 211, 214, 215, 216, 221–222, 230, 232
Chaperones
collagen triple helix and, 326
elastin and, 443, 444, 448
stress proteins/IFAPs and, 173–174
Charge-charge complementarity
in coiled coil design, 91
Charged residues
sequence repeats and, 20
Charged/apolar residue ratio
IF chains and, 124
Circular dichroism (CD) studies
elastin and, 447–448

Cirrhosis
fibrinogen/fibrin and, 256
Clotting
fibrinogen/fibrin and, 248, 250, 256–257, 263, 264, 268–269, 270, 272–273, 277–278, 280
Coacervation
in elastin, 444–445
Coassembly
collagen fibrils and, 344, 345
Coiled coil design
ABC design in, 95
α-helix/amphipathic and, 82–83
antiparallel structures in, 101–105
APC tumor suppressor protein probe and, 96
β-sheets and, 83
belt-and-braces design and, 99–100
buried polar groups and, 91
canonical/non-canonical, 82–83, 91
charge-charge complementarity in, 91
coiled-coil motifs in, 81–82
coil-Ser design in, 93, 94
conformational state switch in, 99–100
Fos:Jun system and, 91, 94
globular proteins and, 83
helical wheels and, 84, 85
heptad repeats and, 82–83, 84, 93
heterotetramer design in, 95–96
Hodges design in, 92–93, 102–103
homodimer design in, 104–105
hydrophobicity and, 83–86, 87, 89, 90
leucine zipper and, 85, 87–88, 91, 97, 101, 104
oligomer specification and, 86–90
packing geometries in, 86–90
packing/acute in, 86, 89
packing/parallel in, 86, 88
parallel structures in, 92–101
PDB and, 86–90
Peptide Velcro design in, 93–95, 103–104
polar residues and, 89, 91, 92
preferred repeats and, 84, 86
probes and, 79
protein engineering and, 80
protein redesign and, 80
right-handed coils in, 96–97
SAF design in, 97–99
salt-bridges in, 90–91, 99
saturation v. random mutagenesis and, 81
sequence-to-structure rules in, 81, 82–92
shortest design and, 80, 99
spectroscopic analysis in, 101
stability in, 90–91
structure/function and, 79, 80
x-ray analysis and, 87, 98
Coiled coil structures
α-helix form of, 38–39
apolipoproteins and, 67, 69, 70
compound helices of, 39
Crick's model and, 39, 40, 43–44, 55
crossing angle in, 70
deviations in, 53–56
discontinuities of, 49, 50, 52–53
epidermin and, 38
fibrinogen and, 71–72
folding/stability of, 56–60
functions of, 70–72
funnel structures and, 67, 68
heptads and, 45, 51, 53, 57, 61
hourglass structures and, 63, 66, 67
hydrophobicity and, 40, 42, 43, 53, 54, 56–57, 59, 60
IFs in, 72
insertions/periodicity in, 49, 50, 51
keratin and, 38, 45, 57–58, 71–72
knobs-into-holes and, 39, 40, 42, 48, 49, 51, 53, 54–55, 63, 67, 70
knobs-to-knobs and, 54–56, 70
leucine zippers and, 42, 45, 56, 57, 58, 61, 70, 72
MexA and, 63, 66, 67
myosin/tropomyosin and, 38, 45, 57–58, 72
oligomer specificity and, 57–58, 61
PDB and, 48, 61
pitch angle of, 44–45
prediction/analysis programs and, 45–49
ridges-into-grooves and, 40, 61, 70
Sendai virus and, 53, 60
sheet proteins and, 63, 65, 67
spectrin and, 40, 42
stammers/stutters/skips and, 49, 54, 55
standard model and, 40–45
structural diversity of, 60–70
structural parameters of, 46
supercoil axis and, 40, 43–44
superhelical distortion of, 39
tetrabrachion stalk and, 52, 53–54, 60

Coiled coil structures (*cont.*)
trigger sequence and, 59
tube structures and, 63, 64
x-ray diffraction of, 37–38, 60
zinc fingers and, 61
Coiled-coil motifs
in coiled coil design, 81–82
Coiled-coil proteins
fibrinogen/fibrin and, 256, 258, 259
fibrous proteins and, 1, 2, 3–4, 5, 8
Coil-Ser design
in coiled coil design, 93, 94
Collagen fibrils
AGP and, 359
axial packing in, 345–346, 347
coassembly and, 344, 345
collagen triple helix and, 303–304, 306, 307, 315, 318, 330
crosslinks and, 343, 346, 349, 350
diameter of, 341, 342, 343, 344, 358, 359, 360
D-repeat periodicity in, 341, 342, 344, 346, 350, 354, 356, 363, 366
EM analysis of, 344, 346, 349, 362, 364
FACIT collagens and, 342–343, 345, 357
fibril surfaces in, 356–357
fibril-forming collagens and, 342–343
FLS fibrils and, 355–356
GAGs and, 362, 363
growth regulation of, 357–359
kinks/crimps in, 345, 350, 361, 362, 363
knockout mice studies and, 358–359
lateral packing in, 345–346, 347, 350–354
mechanical properties of, 365–366
models of, 351–354, 356, 363
molecular packing and, 345–350, 351, 356
mutagenesis and, 345
NC4 domains and, 357
NMR analysis of, 347–348, 349
procollagens and, 357–358, 363
proteoglycans and, 341, 342, 357, 358, 361, 362, 363
S-folds in, 349
size distribution of, 359–361
SPARC and, 359
stabilization of, 346
stress/strain and, 365–366
subfibrillar structures of, 356
suprabundle structures and, 365
suprafibrillar structures of, 361–365
telopeptides in, 346–349
tensile strength/creep and, 359, 366–367
tilts in, 350, 353–354
type I collagen-rich fibrils and, 343–344
type II collagen-rich fibrils and, 343–344
x-ray analysis of, 346–348, 349, 350, 351, 365, 366
Collagen triple helix
chaperones and, 326
collectins and, 306
domains of, 305, 306–307, 308, 328, 329
EKG peptide and, 308, 309, 315, 316, 318
EM of, 326–327
endostatin and, 303–304
fiber diffraction analysis of, 302, 307, 308
fibrils and, 303–304, 306, 307, 315, 318, 330
ficolins and, 306
Flp and, 322–323, 332
Gly-X-Y repeats and, 301–302, 311–312, 319, 320, 323, 327, 329, 330
helical twists in, 308, 311
hydration networks and, 314–315, 316, 322
hydrogen bonding in, 308, 310, 311–312, 313, 322
hydrophobicity in, 315–316
IBP and, 308, 309, 315, 316, 327, 328, 329, 332
imidic acids and, 301–302, 308, 312, 318, 319, 322
ligand binding and, 307, 326–328
MASPs and, 326, 327, 328
MBL and, 326, 327, 330, 332
molecular dynamics studies in, 318–319, 331
molecular packing in, 316–318
molecular structure of, 307–319
motifs role in, 303–307
mutations/disease and, 301–302, 329–332
NMR analysis and, 302–303, 307, 312, 314, 316, 318, 332
nucleation and, 307, 311
OI and, 329, 331, 332
PDB and, 319
proteoglycans and, 326
puckers in, 321, 322, 323
recombinant studies and, 316, 319, 324, 327
sequence repeats and, 301–302
side chain reactions in, 310, 312, 315

stability of, 302, 323–326
stabilization in, 317, 319–323
staggering in, 302, 306, 317, 318, 324
structural role of, 303–304
water and, 310, 312, 315
x-ray analysis and, 302–303, 307, 311, 331
Collagen type IV
network-forming collagens and, 377–380
Collagen type VI
network-forming collagens and, 380–386
Collagen type VIII
network-forming collagens and, 386–387
Collagen type X
network-forming collagens and, 387–388
Collagens, 1, 9, 10, 12. *See also* Collagen triple helix; Collagens, network-forming; Fibrous proteins
crystal structure of, 7
FACIT collagen and, 7–8
fibrils of, 1, 6, 7
sequence repeats and, 13, 17–18
types of, 7, 8
Collagens, network-forming
aging and, 376
Alport syndrome and, 379
assembly of, 379–380, 384, 385, 389
basement membranes and, 376, 379, 386
beaded filaments and, 376
Bethlem myopathy and, 376
crosslinking in, 377, 379–380
Descemet's membrane and, 386–387
disease and, 376
dogfish egg case model and, 375, 376, 388–392, 398
EM studies of, 378, 380, 388
fibers v., 375–376
flexibility/stiffness of, 377, 399
globular domains in, 377, 380–383, 392, 396, 398
Gly-X-Y repeats and, 377, 386, 388, 389, 392, 399
Golgi and, 389
hexagonal aggregation and, 376, 389, 390, 391
hydrophobic regions in, 380, 386, 388, 391–392, 395, 396
integrins and, 387
interruptions in, 377, 399
linear/lateral aggregation and, 376, 380, 388, 396, 399
nodules in, 388
Sorsby's fundus dystrophy in, 382–384, 385
stability in, 379, 381–382
structural themes in, 396–399
supercoiling in, 377, 382, 387–388, 392–396, 398, 399
TIMP3 and, 384, 385
type IV collagen and, 377–380
type VI collagen and, 380–386
type VIII collagen and, 386–387
type X collagen and, 387–388
Ullrich syndrome and, 376
von Willebrand domains in, 380–381
x-ray analysis of, 388
Collectins
collagen triple helix and, 306
Conformational state switch
in coiled coil design, 99–100
Connective tissue
fibrous proteins and, 6–9
Contiguous repeats
sequence repeats and, 12, 19
Costameres
IFAP structure/function and, 164–165, 181
spectrin superfamily and, 213
Crick's model
coiled coil structures and, 39, 40, 43–44, 55
Crossing angle
coiled coil structures and, 70
Crosslinking
collagen fibrils and, 343, 346, 349, 350
in elastin, 437, 439, 441, 442, 445–446, 447, 453
fibrillin microfibrils and, 415, 417, 420, 423
fibrinogen/fibrin and, 247, 257, 266, 270–272, 284
IF chains and, 124, 127, 128
IFAP structure/function and, 144, 146, 170, 171
in network-forming collagens, 377, 379–380
spectrin superfamily and, 204, 207, 209, 211–212, 222
Cytokines
fibrinogen/fibrin and, 261

Cytoskeletal proteins
spectrin superfamily and, 204, 208, 212–213, 220, 221, 224, 225, 230, 232

D

Data analysis
coiled coil design and, 86, 89–90
coiled coil structures and, 47, 48, 61
collagen triple helix and, 319
fibrous proteins and, 13
PROSITE database and, 13
Deletions/insertions
IF chains and, 117, 118
Descemet's membrane
network-forming collagens and, 386–387
Desialylation
fibrinogen/fibrin and, 256
Desmocalmin/keratocalmin
IFAP structure/function and, 161
Desmoplakins
IFAP structure/function and, 149–152
Desmosomes
IFAP structure/function and, 146, 147, 149, 158
Desmoyokin
sequence repeats and, 19
Diabetes
fibrinogen/fibrin and, 283
Diameters
of collagen fibrils, 341, 342, 343, 344, 358, 359, 360
Disorders
of elastin, 450–451
DMD. *See* Duchenne muscular dystrophy
Dogfish egg case model
network-forming collagens and, 375, 376, 388–392, 398
Domains
of collagen triple helix, 305, 306–307, 308, 328, 329
fibrillin microfibrils and, 406–407
of fibrinogen/fibrin, 253–256, 264, 265, 270, 272
DP. *See* Desmoplakins
DPC. *See* Dystrophin-associated protein complex
D-repeat periodicity
collagen fibrils and, 341, 342, 344, 346, 350, 354, 356, 363, 366
DRP. *See* Dystrophin-related protein
Duchenne muscular dystrophy (DMD)
spectrin superfamily and, 212–213, 219, 227–229
Dysfibrinogenemias
fibrinogen/fibrin and, 264, 279–283, 284
Dystrobrevin
IFAP structure/function and, 164, 165–166
Dystrophin-associated protein complex (DPC)
IFAP structure/function and, 162, 164, 165–166
Dystrophin-related protein (DRP)
spectrin superfamily and, 213
Dystrophins
fibrous proteins/MS, 5, 10
spectrin superfamily and, 205, 208, 212–213, 215, 217, 226, 227

E

EBP. *See* Elastin-binding protein chaperone
EF hands. *See* Calcium binding/EF hands
EKG peptide
collagen triple helix and, 308, 309, 315, 316, 318
Elastic fibers
fibrous proteins and, 1, 5, 6, 8
Elasticity
of elastin, 437, 448, 449–450
of fibrillin microfibrils, 405–406, 419–420, 423
Elastin
aging and, 451
allysine and, 445–446
assembly of, 444–445
biological properties of, 448
as biomaterial, 448, 450–453
cbEGF repeats, 439
CD studies of, 447–448
coacervation in, 444–445
crosslinking in, 437, 439, 441, 442, 445–446, 447, 453
disorders of, 450–451
EBP chaperone and, 443, 444, 448

elasticity/flexibility of, 437, 448, 449–450
elastogenesis and, 437–438, 440–446, 453
ELR and, 448
EM studies of, 447–448
expression in, 441–443
fibrillar model of, 449–450
fibrous proteins and, 1, 2, 9, 10
Golgi and, 443
hydrogels and, 452
hydrophobic domains in, 444–445, 447
liquid drop model of, 449
LTBPs and, 439–440
MAGPs and, 439, 444
MFAPs and, 440
microfibrils and, 439–440
oiled coil model of, 449–450
physical properties of, 445–446
proteoglycans and, 440
random chain model of, 449
secretion of, 443
sources of, 438–439
stress/strain in, 446
structural properties of, 447–448
structure of, 438–439
synthetic matrices of, 452–453
TGFβ and, 443
transcription in, 441–442
tropoelastin and, 437, 439, 440, 443, 444, 445
Elastin-binding protein (EBP) chaperone
elastin and, 443, 444, 448
Elastin-laminin receptor (ELR)
elastin and, 448
Elastogenesis
elastin and, 437–438, 440–446, 453
Electron microscope (EM) analysis
of elastin, 447–448
Electron microscopy (EM) analysis
of collagen fibrils, 344, 346, 349, 362, 364
collagen triple helix and, 326–327
of fibrillin microfibrils, 417
fibrinogen/fibrin and, 251, 254, 265, 268
of network-forming collagens, 378, 380, 388
sequence repeats and, 25
ELR. *See* Elastin-laminin receptor
EM. *See* Electron microscopy analysis
Endostatin
collagen triple helix and, 303–304
Envoplakin/periplakin
IFAP structure/function and, 157–158
Epiplakin
IFAP structure/function and, 149
sequence repeats and, 13
Evolution
of fibrinogen/fibrin, 283–284
Expression
of elastin, 441–443
Extracellular matrix assemblies
fibrillin microfibrils and, 405

F

FACIT collagens
collagen fibrils and, 342–343, 345, 357
collagens and, 7–8
Factor XIIIa
fibrinogen/fibrin and, 248, 257–258, 263, 266, 270–271, 274, 276, 284
Fiber diffraction analysis
of collagen triple helix, 302, 307, 308
Fibril surfaces
collagen fibrils and, 356–357
Fibril-forming collagens
collagen fibrils and, 342–343
Fibrillar model
of elastin, 449–450
Fibrillin
fibrous proteins and, 1, 8, 9, 10
Fibrillin microfibrils
AFM and, 411, 417, 420
aging and, 426
assembly of, 412–417
beads-on-a-string and, 405, 406, 407
calcium chelation in, 419
cbEGF domains in, 407, 408–409, 410, 411, 415, 420, 427, 428
crosslinking in, 415, 417, 420, 423
domains of, 406–407
elasticity of, 405–406, 419–420, 423
EM of, 417
as extracellular matrix assemblies, 405
fibrillin-1 homotypic interactions and, 412–414
fibrillinopathies of, 426–429
fibrillogenesis and, 406
flexibility in, 409, 411
folding in, 412

Fibrillin microfibrils (*cont.*)
glycosylation of, 406–407, 411, 412, 413
hinged alignment model in, 420–423
HSPGs and, 412
hybrid motifs and, 410
hydrophilicity and, 411
hydrophobicity and, 410, 411
integrins and, 409–410, 412
linkers and, 409
LTBPs and, 415–416
MAGP interactions and, 414–416, 421–422
mass spectroscopy studies of, 411
maturation of, 417, 420
MFAP interactions and, 414–416
MFS and, 405, 410, 427–429
molecular organization of, 411
mutations of, 406, 409, 427–429
N-/C- termini and, 410–411, 413, 414, 423
NMR studies of, 408
PDI and, 411
periodicity and, 419
proline/glycine-rich regions and, 410
proteoglycans and, 412, 416, 426
recombinant studies of, 413
RGD integrin recognition site and, 409–410, 415
staggered alignment model in, 420–423, 424–426
staggering in, 419, 420
STEM mass mapping and, 417, 418, 419, 423
TAMD and, 422, 425
tandem repeats of, 408
TB domains in, 407
TB motifs in, 407–408, 409–410, 420
TGFβ and, 409–410, 415–416, 428–429
tropoelastin and, 415, 416, 421–422
x-ray analysis of, 408, 419
Fibrillin-1 homotypic interactions
fibrillin microfibrils and, 412–414
Fibrillinopathies
of fibrillin microfibrils, 426–429
Fibrillogenesis
fibrillin microfibrils and, 406
Fibrils, type I collagen-rich. *See also* Collagen fibrils
collagen fibrils and, 343–344
Fibrils, type II collagen-rich
collagen fibrils and, 343–344
Fibrin polymerization
fibrinogen/fibrin and, 248–249, 263, 264–269, 270
Fibrin/fibrinogen
fibrous proteins and, 5–6
Fibrinogen
coiled coil structures and, 71–72
sequence repeats and, 20–21
Fibrinogen/fibrin
afibrinogenemias in, 281
albumin and, 273–274
binding pockets and, 269
biosynthesis of, 260–262
calcium binding and, 247, 257, 259–260, 269, 270, 277, 284
carcinomas and, 261, 282, 283
cell adhesion and, 274, 278, 285
chain assembly in, 261–262
cirrhosis, 256
clotting and, 248, 250, 256–257, 263, 264, 268–269, 270, 272–273, 277–278, 280
coiled coils and, 256, 258, 259
crosslinking in, 247, 257, 266, 270–272, 284
cytokines and, 261
desialylation of, 256
diabetes and, 283
domains of, 253–256, 264, 265, 270, 272
dysfibrinogenemias in, 264, 279–283, 284
EM analysis of, 251, 254, 265, 268
evolution of, 283–284
factor XIIIa and, 248, 257–258, 263, 266, 270–271, 274, 276, 284
fibrin polymerization and, 248–249, 263, 264–269, 270
fibrinolysis and, 248–250, 264, 275–277
fibrinopeptides and, 247–248, 251, 253, 260, 263–264, 267, 270, 278–279, 280–281
fibroblast growth factor-2 and, 274–275
fibroblasts and, 277
fibronectin and, 271, 273
fibulin and, 274
glycoprotein structure of, 247, 248
glycosylation of, 256–257
hydrophobicity and, 256, 259, 263
integrins and, 248, 273–274, 277–279, 285
interleukins and, 274–275
knobs/holes in, 248, 254, 257, 258, 264, 284

leukocytes and, 279
Lp(a) and, 276–277
microcalorimetry of, 254–255
mutagenesis of, 260, 264, 279–281
physico-chemical properties of, 251–253, 259
plasminogen and, 248–250, 275, 277
platelets and, 249–250, 262, 274, 277–279, 281, 285
protofibrils and, 266, 267–268, 273
signaling and, 260–261
sources of, 262
structure/properties of, 251–260
TAFI and, 274–275
thrombin and, 247, 258, 263–264, 277
thrombospondin and, 274
tPA and, 275–277
transgenic mice and, 278, 279
vascular endothelial growth factor and, 274–275
von Willebrand factor and, 274
x-ray analysis of, 248, 251, 255, 256, 257–259, 265, 270, 284
Fibrinolysis
fibrinogen/fibrin and, 248–250, 264, 275–277
Fibrinopeptides
fibrinogen/fibrin and, 247–248, 251, 253, 260, 263–264, 267, 270, 278–279, 280–281
Fibroblast growth factor-2
fibrinogen/fibrin and, 274–275
Fibroblasts
fibrinogen/fibrin and, 277
Fibronectin
fibrinogen/fibrin and, 271, 273
Fibrous long spacing (FLS) fibrils
collagen fibrils and, 355–356
Fibrous proteins, 1, 2. *See also* Collagens; Sequence repeats, in fibrous proteins
actinins and, 1, 5, 10
amyloids and, 2
coil-coiled proteins and, 1, 2, 3–4, 5, 8
collagens and, 1, 2, 6–8
connective tissue and, 6–9
dystrophin/MS and, 5, 10
EF hands and, 5, 27–28
elastic fibers in, 1, 5, 6, 8
elastins and, 1, 2, 9, 10
fibrillin, 1, 8, 9, 10
fibrin/fibrinogen and, 1, 2, 5–6
α-fibrous proteins and, 2–6, 10, 12
filaments and, 1, 2
IFAPs and, 4–5
IFs and, 4–5
keratins and, 4
NMR and, 2, 5, 9
prions and, 2
proteoglycans and, 6
repetitive motifs/heptads and, 2–3
spectrins and, 1, 5, 10
tropoelastin and, 8–9
tropomyosin and, 3
x-ray crystallography of, 2, 5–6, 9
α-fibrous proteins
fibrous proteins and, 2–6, 10
sequence repeats and, 14
Fibulin
fibrinogen/fibrin and, 274
Filaggrin
IFAP structure/function and, 168–169
sequence repeats and, 17
Filament formation
fibrous proteins and, 12
sequence repeats and, 12, 20–21
Filaments, beaded
network-forming collagens and, 376
Fimbrin
IFAP structure/function and, 172
Flexibility
of elastin, 437, 448, 449–450
in fibrillin microfibrils, 409, 411
of IF chains, 124–125, 127–128
of network-forming collagens, 377, 399
spectrin superfamily and, 215, 217, 219
Flp. *See* Fluoroproline
FLS. *See* Fibrous long spacing fibrils
Fluoroproline (Flp)
collagen triple helix and, 322–323, 332
Focal contacts
IFAP structure/function and, 162–163
spectrin superfamily and, 211
Focal segmental glomerulosclerosis (FSGS)
spectrin superfamily and, 231–232
Folding
coiled coil structures and, 56–60
in fibrillin microfibrils, 412
Fos:Jun system
coiled coil design and, 91, 94

FSGS. *See* Focal segmental glomerulosclerosis
Funnel structures
coiled coil structures and, 67, 68

G

GAG. *See* Glycosaminoglycans
Globular domains
in network-forming collagens, 377, 380–383, 392, 396, 398
Globular proteins
coiled coil design and, 83
Glycine residues
sequence repeats and, 18–19
Glycoprotein structure
of fibrinogen/fibrin, 247, 248
Glycosaminoglycans (GAG)
collagen fibrils and, 362, 363
Glycosylation
of fibrillin microfibrils, 406–407, 411, 412, 413
of fibrinogen/fibrin, 256–257
Gly-X-Y repeats
collagen triple helix and, 301–302, 311–312, 319, 320, 323, 327, 329, 330
in network-forming collagens, 377, 386, 388, 389, 392, 399
Golgi
elastin and, 443
IFAP structure/function and, 177–178
network-forming collagens and, 389
G-protein signaling
IFAP structure/function and, 161–162
Growth regulation
of collagen fibrils, 357–359

H

H1 subdomains
IF chains and, 117, 119, 127
H2 subdomains
IF chains and, 132–133
Head domains
IF chains and, 116–120
Head-to-tail domains
sequence repeats and, 23, 24–25
HEAT repeats
sequence repeats and, 30
Helical twists
in collagen triple helix, 308, 311
Helical wheels
coiled coil design and, 84, 85
Helices, compound
coiled coil structures and, 39
Helix initiation motif
IF chains and, 121
Helix termination motif
IF chains and, 131
Heparan sulfate proteoglycans (HSPG)
fibrillin microfibrils and, 412
Heptads
coiled coil design and, 82–83, 84, 93
coiled coil structures and, 45, 51, 53, 57, 61
fibrous proteins and, 2–3
IF chains and, 114, 120, 121, 124, 127, 128, 129, 130
sequence repeats and, 19–21, 25
Heterotetramer design
coiled coil design and, 95–96
Hexagonal aggregation
network-forming collagens and, 376, 389, 390, 391
Hinged alignment model
in fibrillin microfibrils, 420–423
Hodges design
coiled coil design and, 92–93, 102–103
Homodimer design
coiled coil design and, 104–105
Homopolymerization
IF chains and, 114–115
Hourglass structures
coiled coil structures and, 63, 66, 67
HSPG. *See* Heparan sulfate proteoglycans
Hybrid motifs
fibrillin microfibrils and, 410
Hydration networks
collagen triple helix and, 314–315, 316, 322
Hydrogels
elastin and, 452
Hydrogen bonding
in collagen triple helix, 308, 310, 311–312, 313, 322
Hydrophilicity
fibrillin microfibrils and, 411

Hydrophobic regions
coiled coil design and, 83–86, 87, 89, 90
coiled coil structures and, 40, 42, 43, 53, 54, 56–57, 59, 60
in collagen triple helix, 315–316
in elastin, 444–445, 447
fibrillin microfibrils and, 410, 411
fibrinogen/fibrin and, 256, 259, 263
in network-forming collagens, 380, 386, 388, 391–392, 395, 396

I

IF. *See* Intermediate filaments
IF chains, sequence/structure
acidic residues and, 117, 121, 122, 125, 128, 132
aggregation characteristics and, 122
apolar residues and, 117, 118, 121, 124, 126, 129
assembly of, 131–132, 136
β-sheets and, 120, 133
basic residues and, 117, 121, 122, 125, 126, 132
charged/apolar residue ratio and, 124
crosslinking and, 124, 127, 128
deletions/insertions and, 117, 118
flexibility of, 124–125, 127–128
H1 subdomain and, 117, 119, 127
H2 subdomain and, 132–133
head domains and, 116–120
helix initiation motif and, 121
helix termination motif and, 131
heptad repeats and, 114, 120, 121, 124, 127, 128, 129, 130
homopolymerization in, 114–115
ionic interactions of, 113–114, 121, 126, 128, 131
keratins and, 113, 114, 117–120, 127, 128, 132–135, 137
key residues and, 113–114, 131
linker L1 and, 123–125
linker L2 and, 128–129
linker L12 and, 127–128
mutations and, 122–123, 125, 129
nestins and, 114
NFs and, 114, 135
nucleating point and, 126
posttranslational modifications and, 135–136
rod domains and, 135
segment 1A and, 120–123
segment 1B and, 125–127
segment 2A and, 128
segment 2B and, 129–132
stability of, 123, 126
stammers/stutters/skips and, 124, 129, 131
structural domains and, 115–116
tail domains and, 132–136
trigger motif and, 126, 131
x-ray analysis of, 127
IFAP. *See* Intermediate filament associated proteins
IFAP structure/function, 143–144
14-3-3 proteins and, 176–177
ABD proteins and, 146, 155, 156, 158, 167, 176–177
actin rich cortex and, 162
actins and, 146, 156, 171–173
apoptosis and, 174–175
armadillo proteins and, 159–161
BP230 isoforms and, 152–154
calcium binding/EF hands and, 168, 169, 170
calponin/CH domains and, 153, 155, 156, 171–172
cell cycles/kinases and, 175–176
cell homeostasis and, 144, 181
cell process control, 181–182
cell surface interactions and, 159–168
chaperones/stress proteins and, 173–174
costameres and, 164–165, 181
crosslinking and, 144, 146, 170, 171
desmocalmin/keratocalmin and, 161
desmoplakins and, 149–152
desmosomes and, 146, 147, 149, 158
DPCs and, 162, 164, 165–166
dystrobrevin and, 164, 165–166
envoplakin/periplakin and, 157–158
epiplakin and, 149
filaggrin and, 168–169
fimbrin and, 172
focal contacts and, 162–163
G-protein signaling and, 161–162
IF families and, 144, 145, 146
KAPs and, 170–171
keratin IFs and, 180

IFAP structure/function (*cont.*)
lamins and, 144, 181
membrane trafficking and, 177–178
membranes/Golgi and, 177–178
motor molecules and, 178–180
muscle cell surface and, 163–165
muscular dystrophy and, 165, 166
myosin va and, 179
nebulin and, 172–173
NFs and, 167–168, 175, 179–181
nudel and, 179
Parkinson's disease and, 179
pinin and, 161
PKPs and, 160–161
plakins and, 146–148, 149
plectin and, 154–157
polycystin-1 and, 161–162
PRDs and, 150, 157, 159
sequence specificity and, 150–151
signaling/cell metabolism and, 144, 173–178
spectrin/ankyrin and, 166, 167–168
trichohyalin and, 169–170
vimentin and, 163, 169, 176, 178–179
x-ray analysis and, 150
z-discs and, 163–164, 172–173
Imidic acids
collagen triple helix and, 301–302, 308, 312, 318, 319, 322
Insertions
coiled coil structures and, 49, 50, 51
Integrin binding protein (BP)
collagen triple helix and, 308, 309, 315, 316, 327, 328, 329, 332
Integrins
fibrillin microfibrils and, 409–410, 412
fibrinogen/fibrin and, 248, 273–274, 277–279, 285
network-forming collagens and, 387
Interleukins
fibrinogen/fibrin and, 274–275
Intermediate filament associated proteins (IFAP)
fibrous proteins and, 4–5
Intermediate filaments (IF)
coiled coil structures and, 72
families of, 144, 145, 146
fibrous proteins and, 4–5
Interruptions
in network-forming collagens, 377, 399
Ionic interactions
IF chains and, 113–114, 121, 126, 128, 131

K

KAP. *See* Keratin-associated proteins
Keratin IFs
IFAP structure/function and, 180
Keratin-associated proteins (KAP)
IFAP structure/function and, 170–171
sequence repeats and, 21–22, 23
Keratins
coiled coil structures and, 38, 45, 57–58, 71–72
fibrous proteins and, 4
IF chains and, 113, 114, 117–120, 127, 128, 132–135, 137
sequence repeats and, 14, 16, 17, 31
Key residues
IF chains and, 113–114, 131
Kinks/crimps
collagen fibrils and, 345, 350, 361, 362, 363
Knobs-into-holes
coiled coil structures and, 39, 40, 42, 48, 49, 51, 53, 54–55, 63, 67, 70
fibrinogen/fibrin and, 248, 254, 257, 258, 264, 265, 266, 284
Knobs-to-knobs
coiled coil structures and, 54–56, 70
Knockout mice studies
collagen fibrils and, 358–359

L

Lamins
IFAP structure/function and, 144, 181
Latent TGFβ-binding proteins (LTBP)
elastin and, 439–440
fibrillin microfibrils and, 415–416
Lateral packing
in collagen fibrils, 345–346, 347, 350–354
Leucine zippers
coiled coil design and, 85, 87–88, 91, 97, 101, 104

coiled coil structures and, 42, 45, 56, 57, 58, 61, 70, 72
sequence repeats and, 19–20
Leucine-rich repeats (LRR)
sequence repeats and, 29–30
Leukocytes
fibrinogen/fibrin and, 279
Ligand binding
collagen triple helix and, 307, 326–328
Linear/lateral aggregation
network-forming collagens and, 376, 380, 388, 396, 399
Linkers
fibrillin microfibrils and, 409
IF chains and, 123–125, 127–128, 129
spectrin superfamily and, 209, 212, 215, 218, 220
Lipid binding
spectrin superfamily and, 224–225
Lipoprotein a (Lp(a))
fibrinogen/fibrin and, 276–277
Liquid drop model
of elastin, 449
Lp(a). *See* Lipoprotein a
LRR. *See* Leucine-rich repeats
LTBP. *See* Latent TGFβ-binding proteins

M

MAGP interactions
elastin and, 439, 444
fibrillin microfibrils and, 414–416, 421–422
Mannose binding lectin (MBL)
collagen triple helix and, 326, 327, 330, 332
Marfan syndrome (MFS)
fibrillin microfibrils and, 405, 410, 427–429
fibrous proteins and, 9
MASPs
collagen triple helix and, 326, 327, 328
Maturation
of fibrillin microfibrils, 417, 420
MBL. *See* Mannose binding lectin
Membrane trafficking
IFAP structure/function and, 177–178
MexA
coiled coil structures and, 63, 66, 67
MFAP. *See* Microfibril-associated protein interactions
MFS. *See* Marfan syndrome
Microcalorimetry
of fibrinogen/fibrin, 254–255
Microfibril-associated protein (MFAP) interactions
elastin and, 440
fibrillin microfibrils and, 414–416
Microfibrils
elastin and, 439–440
Models
of collagen fibrils, 351–354, 356, 363
Molecular dynamics studies
collagen triple helix and, 318–319, 331
Molecular packing
collagen fibrils and, 345–350, 351, 356
in collagen triple helix, 316–318
Molecular trafficking
spectrin superfamily and, 204
Motifs
PROSITE database of, 13
repetitive, 2–3
structural/functional repeats and, 11–12, 17, 18
Motor molecules
IFAP structure/function and, 178–180
Muscle cell surfaces
IFAP structure/function and, 163–165
Muscular dystrophy
IFAP structure/function and, 165, 166
Mutations
coiled coil design and, 81
collagen fibrils and, 345
collagen triple helix and, 301–302, 329–332
fibrillin microfibrils and, 406, 409, 427–429
fibrinogen/fibrin and, 260, 264, 279–281
IF chains and, 122–123, 125, 129
Myosin va
IFAP structure/function and, 179
Myosin/paramyosin
sequence repeats and, 20–21, 25
Myosin/tropomyosin
coiled coil structures and, 38, 45, 57–58, 72

N

N-/C- termini
fibrillin microfibrils and, 410–411, 413, 414, 423
NC4 domains
collagen fibrils and, 357
Nebulin
IFAP structure/function and, 172–173
Nestins
IF chains and, 114
sequence repeats and, 19
Neurofilament (NF) chains
IF chains and, 114, 135
IFAP structure/function and, 167–168, 175, 179–181
NF-H sequence repeats and, 15–16, 19
NF. *See* Neurofilament chains
NF-H. *See* Neurofilament (NF) chains
NMR. *See* Nuclear magnetic resonance
NMR analysis
collagen triple helix and, 302–303, 307, 312, 314, 316, 318, 332
Nuclear magnetic resonance (NMR) analysis
of collagen fibrils, 347–348, 349
of fibrillin microfibrils, 408
fibrous proteins and, 2, 5, 9
sequence repeats and, 17, 31
Nucleating point
IF chains and, 126
Nucleation
collagen triple helix and, 307, 311
Nudel
IFAP structure/function and, 179

O

OI. *See* Osteogenesis imperfecta
Oiled coil model
of elastin, 449–450
Oligomer specification
coiled coil design and, 86–90
Oligomer specificity
coiled coil structures and, 57–58, 61
Osteogenesis imperfecta (OI)
collagen triple helix and, 329, 331, 332

P

Packing geometries
coiled coil design and, 86–90
Packing/acute
coiled coil design and, 86, 89
Packing/parallel
coiled coil design and, 86, 88
Parallel structures
coiled coil design and, 92–101
Parkinson's disease
IFAP structure/function and, 179
PDB. *See* Protein Data Bank
PDI. *See* Protein disulphide isomerase
Peptide Velcro design
coiled coil design and, 93–95, 103–104
Periodicity
fibrillin microfibrils and, 419
PH. *See* Pleckstrin homology domain
Phosphorylation/dephosphorylation
sequence repeats and, 16
Pinin
IFAP structure/function and, 161
Pitch angle
coiled coil structures and, 44–45
Pitch length
sequence repeats and, 19–20
PKP. *See* Plakophilins
Plakin repeat domains (PRD)
IFAP structure/function and, 150, 157, 159
Plakins
IFAP structure/function and, 146–148, 149
Plakophilins (PKP)
IFAP structure/function and, 160–161
Plasminogen
fibrinogen/fibrin and, 248–250, 275, 277
Platelets
fibrinogen/fibrin and, 249–250, 262, 274, 277–279, 281, 285
Pleckstrin homology (PH) domain
spectrin superfamily and, 205, 210, 224–225
Plectin
IFAP structure/function and, 154–157
Polar residues
coiled coil design and, 89, 91, 92
Polycystin-1
IFAP structure/function and, 161–162

Polyproline binding domains
spectrin superfamily and, 225–226
Posttranslational modifications
IF chains and, 135–136
PRD. *See* Plakin repeat domains
Prediction/analysis programs
coiled coil structures and, 45–49
Preferred repeats
coiled coil design and, 84, 86
Prions
fibrous proteins and, 2
Procollagens
collagen fibrils and, 357–358, 363
Proline/glycine-rich regions
fibrillin microfibrils and, 410
Protein Data Bank (PDB)
coiled coil design and, 89–90
coiled coil structures and, 48, 61
collagen triple helix and, 319
Protein disulphide isomerase (PDI)
fibrillin microfibrils and, 411
14-3-3 proteins
IFAP structure/function and, 176–177
Proteoglycans
collagen fibrils and, 341, 342, 357, 358, 361, 362, 363
collagen triple helix and, 326
elastin and, 440
fibrillin microfibrils and, 412, 416, 426
fibrous proteins and, 6
sequence repeats and, 29–30
Protofibrils
fibrinogen/fibrin and, 266, 267–268, 273
Puckers
in collagen triple helix, 321, 322, 323

R

Random chain model
of elastin, 449
RBC. *See* Red blood cell studies
Recombinant studies
collagen triple helix and, 316, 319, 324, 327
of fibrillin microfibrils, 413
Red blood cell (RBC) studies
spectrin superfamily and, 210–211, 218, 229–230
Repeats, exact
sequence repeats and, 15, 16
RGD integrin recognition sites
fibrillin microfibrils and, 409–410, 415
Ridges-into-grooves
coiled coil structures and, 40, 61, 70
Right-handed coils
coiled coil design and, 96–97
Rod domains
IF chains and, 135
spectrin superfamily and, 203, 208, 209, 219, 220–221, 228, 231

S

SAF. *See* Self-assembled protein fibers
Salt-bridges
coiled coil design and, 90–91, 99
Scanning transmission EM (STEM) mass mapping
fibrillin microfibrils and, 417, 418, 419, 423
Secreted protein acidic rich in cysteine (SPARC)
collagen fibrils and, 359
Secretion
of elastin, 443
Segment 1A
IF chains and, 120–123
Segment 1B
IF chains and, 125–127
Segment 2B
IF chains and, 129–132
Self-assembled protein fibers (SAF)
coiled coil design and, 97–99
Sendai virus
coiled coil structures and, 53, 60
Sequence repeats
collagen triple helix and, 301–302
Sequence repeats, in fibrous proteins
acidic residues and, 12
actin and, 14
apolar residues and, 12, 15, 19–21
basic residues and, 12
calcium binding/EF hands and, 14–15, 27–28
charged residues and, 20
classes of, 11–15
collagen and, 13, 17–18
contiguous repeats and, 12, 19

Sequence repeats, in fibrous proteins (*cont.*)
desmoyokin and, 19
EM analysis and, 25
epiplakin and, 13
fibrinogen and, 20–21
α-fibrous proteins and, 14
filaggrin/profilaggrin and, 17
filament formation and, 12, 20–21
glycine residues and, 18–19
head-to-tail domains and, 23, 24–25
HEAT repeats and, 30
heptads and, 19–21, 25
KAPS and, 21–22, 23
keratins and, 14, 16, 17, 31
leucine zippers and, 19–20
LRRs and, 29–30
myosin/paramyosin and, 20–21, 25
nestin and, 19
NF-H and, 15–16, 19
NMR and, 17, 31
phosphorylation/dephosphorylation and, 16
pitch length and, 19–20
proteoglycans and, 29–30
SH3 domains and, 22
short/long exact repeats and, 15, 16
structural/functional motifs in, 11–12, 17, 18
stutters and, 21
tropomyosin and, 20–21, 24–25, 26–27
type A, 15–17
type B, 17–19
type C, 19–23
type D, 23–25, 26–27
type E, 27–30
VASP and, 21, 22
ww domains and, 22
x-ray analysis and, 17, 18, 19, 25, 26–27, 31
zinc fingers and, 28–29
Sequence repeats, type A
in fibrous proteins, 15–17
Sequence repeats, type B
in fibrous proteins, 17–19
Sequence repeats, type C
in fibrous proteins, 19–23
Sequence repeats, type D
in fibrous proteins, 23–25, 26–27
Sequence repeats, type E
in fibrous proteins, 27–30
Sequence specificity
IFAP structure/function and, 150–151
Sequence-to-structure rules
coiled coil design and, 81, 82–92
S-folds
in collagen fibrils, 349
SH3 domains
sequence repeats and, 22
spectrin superfamily and, 205, 209, 225–226
Sheet proteins
coiled coil structures and, 63, 65, 67
Side chain reactions
in collagen triple helix, 310, 312, 315
Signaling
fibrinogen/fibrin and, 260–261
IFAP structure/function and, 144, 173–178
spectrin superfamily and, 221–222, 224
Size distribution
of collagen fibrils, 359–361
Sorsby's fundus dystrophy
network-forming collagens and, 382–384, 385
Sources
of elastin, 438–439
of fibrinogen/fibrin, 262
SPARC. *See* Secreted protein acidic rich in cysteine
Spectrin repeats
spectrin superfamily and, 203, 206, 207, 208, 211, 217–219, 220
Spectrin superfamily proteins
ABDs of, 203, 208–210, 213–217, 219, 220, 230
ABS in, 215–217, 222
actinins and, 203, 211, 215, 216, 217, 219, 220–221, 222, 232
actin-stress fibers and, 204, 211
ankyrin and, 211, 220, 225, 230, 231
calcium binding and, 222–224
cell adhesion and, 211
CH domains and, 204, 208, 211, 214, 215, 216, 221–222, 230, 232
costameres and, 213
crosslinking and, 204, 207, 209, 211–212, 222
cytoskeletal proteins and, 204, 208, 212–213, 220, 221, 224, 225, 230, 232
DMD/BMD and, 212–213, 219, 227–229
DRP and, 212–213

dystrophins and, 205, 208, 212–213, 215, 217, 226, 227
EF hands and, 203, 208, 209–210, 222–224, 226
evolution of, 205–208
flexibility and, 215, 217, 219
focal contacts and, 211
FSGS and, 231–232
function of, 210–213
linkers and, 209, 212, 215, 218, 220
lipid binding and, 224–225
members of, 205
molecular trafficking and, 204
PH domain and, 205, 210, 224–225
polyproline binding domains and, 225–226
RBC studies and, 210–211, 218, 229–230
rod domains of, 203, 208, 209, 219, 220–221, 228, 231
SH3 domain and, 205, 209, 225–226
signaling and, 221–222, 224
spectrin repeats and, 203, 206, 207, 208, 211, 217–219, 220
spherocytosis and, 218, 229–231
structure of, 208–210
utrophin and, 205, 206, 208, 210, 213, 215, 217, 220, 224, 226, 227
ww domain and, 205, 210, 224, 226
z-discs and, 211, 224
zz domain and, 210, 226, 227, 228
Spectrins
coiled coil structures and, 40, 42
fibrous proteins and, 1, 5, 10
Spectroscopic analysis
coiled coil design and, 101
Spherocytosis
spectrin superfamily and, 218, 229–231
Stability
in coiled coil design, 90–91
of coiled coil structures, 56–60
of collagen triple helix, 302, 323–326
of IF chains, 123, 126
network-forming collagens and, 379, 381–382
Stabilization
of collagen fibrils, 346
collagen triple helix and, 317, 319–323
Staggered alignment model
in fibrillin microfibrils, 420–423, 424–426
Staggering
in collagen triple helix, 302, 306, 317, 318, 324
in fibrillin microfibrils, 419, 420
Stammers/stutters/skips
coiled coil structures and, 49, 51, 54, 55
IF chains and, 124, 129, 131
sequence repeats and, 21
Standard model
of coiled coil structures, 40–45
STEM. *See* Scanning transmission EM mass mapping
Stress/strain
collagen fibrils and, 365–366
elastin and, 446
Structural discontinuities
coiled coil structures and, 49, 50, 52–53, 53–56
Structural diversity
coiled coil structures and, 60–70
Structural domains
IF chains and, 115–116
Structural properties
of collagen triple helix, 303–304
of elastin, 447–448
of fibrinogen/fibrin, 251–260
Structural themes
in network-forming collagens, 396–399
Structure
of elastin, 438–439
of spectrin superfamily proteins, 208–210
Subfibrillar structures
of collagen fibrils, 356
Supercoil axis
coiled coil structures and, 40, 43–44
Supercoiling
in network-forming collagens, 377, 382, 387–388, 392–396, 398, 399
Superhelical distortion
coiled coil structures and, 39
Suprabundle structures
collagen fibrils and, 365
Suprafibrillar structures
collagen fibrils and, 361–365
Synthetic matrices
of elastin, 452–453

T

TAFI. *See* Thrombin activable fibrinolysis inhibitor
Tail domains
IF chains and, 132–136
TAMD. *See* Theoretical axial mass distributions
Tandem repeats
in fibrillin microfibrils, 408
TB motifs
in fibrillin microfibrils, 407–408, 409–410, 420
Telopeptides
in collagen fibrils, 346–349
Tensile strength/creep
collagen fibrils and, 359, 366–367
Tetrabrachion stalks
coiled coil structures and, 52, 53–54, 60
TGFβ. *See* Transforming growth factor
Theoretical axial mass distributions (TAMD)
fibrillin microfibrils and, 422, 425
Thrombin
fibrinogen/fibrin and, 247, 258, 263–264, 277
Thrombin activable fibrinolysis inhibitor (TAFI)
fibrinogen/fibrin and, 274–275
Thrombospondin
fibrinogen/fibrin and, 274
Tilts
in collagen fibrils, 350, 353–354
TIMP3. *See* Tissue inhibitor metalloproteinase 3
Tissue inhibitor metalloproteinase 3 (TIMP3)
network-forming collagens and, 384, 385
Tissue-type plasminogen activator (tPA)
fibrinogen/fibrin and, 275–277
tPA. *See* Tissue-type plasminogen activator
Transcription
of elastin, 441–442
Transforming growth factor (TGFβ)
elastin and, 443
fibrillin microfibrils and, 409–410, 415–416, 428–429
Transgenic mice
fibrinogen/fibrin studies and, 278, 279
Trichohyalin
IFAP structure/function and, 169–170
Trigger motif
IF chains and, 126, 131
Trigger sequence
coiled coil structures and, 59
Tropoelastin
elastin and, 437, 439, 440, 443, 444, 445
fibrillin microfibrils and, 415, 416, 421–422
fibrous proteins and, 8–9
Tropomyosin
fibrous proteins and, 3
sequence repeats and, 20–21, 24–25, 26–27
Tube structures
coiled coil structures and, 63, 64

U

Ullrich syndrome
network-forming collagens and, 376
Utrophin
spectrin superfamily and, 205, 206, 208, 210, 213, 215, 217, 220, 224, 226, 227

V

Vascular endothelial growth factor
fibrinogen/fibrin and, 274–275
Vasodilator-stimulated phosphoprotein (VASP)
sequence repeats and, 21, 22
VASP. *See* Vasodilator-stimulated phosphoprotein
Vimentin
IFAP structure/function and, 163, 169, 176, 178–179
von Willebrand domains
fibrinogen/fibrin and, 274
in network-forming collagens, 380–381

W

Water
collagen triple helix and, 310, 312, 315

WW domains
sequence repeats and, 22
spectrin superfamily and, 205, 210, 224, 226

X

X-ray analysis
coiled coil design and, 87, 98
coiled coil structures and, 37–38, 60
of collagen fibrils, 346–348, 349, 350, 351, 365, 366
collagen triple helix and, 302–303, 307, 311, 331
of fibrillin microfibrils, 408, 419
of fibrinogen/fibrin, 248, 251, 255, 256, 257–259, 265, 270, 284
of fibrous proteins, 2, 5–6, 9
of IF chains, 127
IFAP structure/function and, 150
of network- forming collagens, 388
sequence repeats and, 17, 18, 19, 25, 26–27, 31

Z

Z-discs
IFAP structure/function and, 163–164, 172–173
spectrin superfamily and, 211, 224
Zinc fingers
coiled coil structures and, 61
sequence repeats and, 28–29
ZZ domains
spectrin superfamily and, 210, 226, 227, 228